INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI.

18891

I. A. R. I. 6.
MGIPC—S4—10 AR—21-6-49—1,000.

PROCEEDINGS

OF THE

ROYAL SOCIETY OF LONDON

SERIES A. MATHEMATICAL AND PHYSICAL SCIENCES

VOL 176

LONDON
Printed and published for the Royal Society
By the Cambridge University Press
Bentley House, N.W.1

27 November 1940

18891
IARI

18891

PRINTED IN GREAT BRITAIN BY
WALTER LEWIS, M.A.
AT THE CAMBRIDGE UNIVERSITY PRESS

CONTENTS

SERIES A VOL 176

No. A 964 28 August 1940

PAGE

New studies on active nitrogen. I Brightness of the after-glow under varied conditions of concentration and temperature By Lord Rayleigh, F.R.S. 1

New studies on active nitrogen. II. Incandescence of metals in active nitrogen, and quantitative estimates of the energy liberated By Lord Rayleigh, F.R.S. 16

X-ray electrons expelled from metals by silver $K\alpha_1$ radiations. By C. J Birkett Clews and H. R. Robinson, F.R S. 28

The electric strength of some solid dielectrics. By A. E W. Austen and S. Whitehead 33

Melting and crystal structure. By J W. H. Oldham and A R Ubbelohde 50

Isomorphous replacement and superlattice structures in the plagioclase felspars. By S. H. Chao and W. H. Taylor 76

The base catalysed prototropy of substituted acetones. By R. P. Bell and O M. Lidwell 88

Potential energy curves in proton transfer reactions. By R. P. Bell and O. M. Lidwell 114

General bi-harmonic analysis for a plate containing circular holes. By A. E. Green 121

Two-phase equilibrium in binary and ternary systems. II. The system methane-ethylene. III. The system methane-ethane-ethylene. By M. Guter, D. M. Newitt and M. Ruhemann 140

No. A 965. 9 October 1940

Bakerian Lecture. Stereochemical types and valency groups. By N. V. Sidgwick, F.R.S. and H. M. Powell 153

The wave form of atmospherics at night. By B. F. J. Schonland, F.R.S., J. S. Elder, D. B. Hodges, W E. Phillips, and J. W. van Wyk. (Plates 1–3) 180

PAGE

A theoretical study of a possible model of paramagnetic alums at low temperatures. By J A Sauer and H. N. V. Temperley . . 203

Operator calculus in the electron theory of metals. By K. Fuchs . . 214

The soft X-ray spectroscopy of solids. II. Emission spectra from simple chemical compounds. By H. M. O'Bryan and H. W. B. Skinner . 229

The distribution of autelectronic emission from single crystal metal points. I. Tungsten, molybdenum, nickel in the clean state By M Benjamin and R. O. Jenkins. (Plates 4–9) 262

Two-phase equilibrium in binary and ternary systems. IV The thermodynamic properties of propane. By J H. Burgoyne . . . 280

No A 966. 1 November 1940

The quenching of the resonance radiation of sodium. By R G W. Norrish, F.R.S , and W. MacF Smith 295

On maintained convective motion in a fluid heated from below. By A. Pellew and R V. Southwell, F R.S 312

Thermal ionization of strontium. By B N. Srivastava . . . 343

The kinetics of mutarotation in solution By J. C. Kendrew and E. A. Moelwyn-Hughes 352

Inner excited states of the proton and neutron. By W. Heitler and S. T. Ma 368

X-ray structure and elastic strains in copper By S L Smith and W. A. Wood. (Plates 10–13) 398

The influence of the solvent on the formation of micelles in colloidal electrolytes I Electrical conductivities of sodium dodecyl sulphate in ethyl alcohol-water mixtures By A. F. H. Ward . 412

No A 967. 27 November 1940

The surface as a limiting factor in the slow combustion of hydrocarbons. By R. G. W. Norrish, F R S , and J. D. Reagh . . . 429

On the turbulence of a tidal current. By J. Proudman, F.R S. . . 449

PAGE

The kinetics of the thermal decomposition of acetophenone. By R. E. Smith and C. N. Hinshelwood, F.R.S. 458

The poisoning of a palladium catalyst by carbon monoxide. By M. G. T. Burrows and W. H. Stockmayer 474

Investigations on the vibration spectra of certain condensed gases at the temperature of liquid nitrogen I Experimental methods A. Infra-red spectra. By G. K. T. Conn, E. Lee and G. B. B. M. Sutherland. B. Raman spectra. By G B B M. Sutherland and Cheng-Kai Wu 484

Investigations on the vibration spectra of certain condensed gases at the temperature of liquid nitrogen II The infra-red and Raman spectra of hydro- and deuterochloric acid at liquid nitrogen temperatures. By E. Lee, G. B. B. M Sutherland and Cheng-Kai Wu . . 493

The flame spectrum of carbon monoxide. By A G. Gaydon (Plate 14) 505

The high-frequency resistance of superconducting tin. By H. London . 522

On the structure of multilayers. I By G Knott, J. H. Schulman and A. F. Wells. (Plate 15) 534

Index 543

New studies on active nitrogen

I. Brightness of the after-glow under varied conditions of concentration and temperature

28. OCT.
LIBRA

By Lord Rayleigh, F.R.S.

(*Received* 3 *May* 1940)

This paper examines quantitatively the behaviour of nitrogen gas emitting the after-glow, under varied conditions.

The integrated light emitted under the most favourable conditions has been measured as 3·18 candle-sec. per c.c of nitrogen, the (rarefied) nitrogen being reckoned as at N.T.P. The number of quanta emitted is estimated as about $1\cdot 3 \times 10^{-3}$ of the number of molecules present.

It is concluded, in agreement with Kneser, that the addition of unexcited nitrogen to glowing nitrogen increases the (instantaneous) emission. It was found that a fivefold increase of total pressure produced about a fivefold increase of brightness. These tests were carried out at very low luminous intensities, when the spontaneous decay was negligible during an experiment

If the active gas contained in 1 litre of weakly glowing nitrogen is allowed to diffuse quickly into an additional 1 litre volume of nitrogen, so as to dilute it by half, the candle power per unit volume is reduced about 4·3-fold. This is a fair approximation to the value 4, which would be expected if the reaction were bimolecular as regards active nitrogen, and is in agreement with the conclusion which has generally been drawn from observation on the rate of decay of the luminosity in closed vessels.

The effect of compressing the glowing gas has been re-examined, using a solid piston moving in a cylinder. It is now found that with the improved arrangements the brightness varies as the inverse cube of the volume. This is the logical conclusion from, and confirmation of, the previous experiments in which the concentration of (1) the inert nitrogen, and (2) the active nitrogen, are separately varied.

A further test was to expand the glowing gas into a doubled volume which was carried out by letting it pass into a supplementary exhausted vessel. This was expected from the previous result to reduce the intensity 8 times, but in fact did not reduce it more than 7 times The gas was therefore somewhat brighter after the expansion than had been expected. Most sources of error would have the opposite effect, and there is an outstanding discrepancy in this result

The effect of temperature on the nitrogen after-glow is examined quantitatively, maintaining the cooler and hotter portion in pressure equilibrium. Cooling to liquid air temperature, for instance, increases the brightness some 80 times. Most, but not all, of this is due to the additional concentration. Assuming (in accordance with the compression experiment) that the brightness is as the cube of the concentration, and correcting the results to uniform concentration, it is found that over the range examined the brightness varies as $T^{-0\cdot 64}$, when T is the absolute temperature.

The positive temperature effect on ordinary chemical reaction corresponds to something more like T^{+100}.

1. NUMBER OF LIGHT QUANTA EMITTED RELATIVE TO NUMBER OF MOLECULES PRESENT

Whatever view may be taken of the processes preceding the emission of light by active nitrogen, the final emission is without doubt by excited molecules. It appeared of interest to determine how the number of excited molecules emitting light would compare with the total number of nitrogen molecules present, when the after-glow was as bright and as persistent as it can easily be made. The number of emitting molecules is taken to be the same as the number of light quanta given out in the whole course of decay.

In order to determine this number by any photometric process, it is in strictness necessary to take into account that they are not all of the same frequency. The after-glow spectrum consists chiefly of three heads of the first positive nitrogen bands. The maximum luminosity of the visual spectrum is at λ 5550, and the wave-lengths of the bands in question and their luminosities in terms of this maximum are as follows (Walsh 1926, p. 471):

Colour	Wave-length	Luminosity
Green	5371	0·930
Yellow	5802	0·870
Red	6251	0·321

The question may be raised whether the α system of bands have great energy in the infra-red. The photograph by Kichlu and Acharya (1929) on a neocyanine plate show the infra-red bands in the region 7500–9000 very feebly compared with any of the visual bands 5053, 5371, and 5755, in spite of the fact that the plate is sensitized for the former and not for the latter. We may fairly infer that the infra-red contribution is not important.

If the light emitted were all of the wave-length 5560, which is the point of maximum luminosity in the spectrum for a normal eye, we could express it in terms of energy by using the "mechanical equivalent of light" $1{\cdot}6 \times 10^{3}$ W per lumen (Ives 1924).

Since the wave-lengths actually in question are of less than the maximum luminosity, the measured emission of light will lead to an underestimate of the energy, and hence of the number of quanta reckoned on the basis of λ 5560. Moreover, the quantum energy of these bands (with the exception of λ 5371) is less than the quantum energy of λ 5560, so this again will probably tend to an underestimate. Nevertheless, the corrections which should be introduced are probably not very important, and could not be found without a difficult spectrophotometric analysis of the light. They will therefore in the first instance be ignored.

It is convenient to calculate as a standard datum the total emission in candle-seconds to be obtained from 1 c.c. of excited nitrogen, at N.T.P., supposing that each molecule gave out one quantum of light $\lambda 5560$, wave number 18,000 cm.$^{-1}$.

The quantum $h\nu = 6{\cdot}55 \times 10^{-27} \times 3 \times 10^{10} \times 18{,}000$

$= 3{\cdot}573 \times 10^{-12}$ erg.

Multiplying this by the number of molecules in c.c., $2{\cdot}75 \times 10^{19}$, we get for the total emission

$$9{\cdot}73 \times 10^{7} \text{ erg.}$$

Now the mechanical equivalent of light is

$$1{\cdot}6 \times 10^{-3} \text{ W per lumen,}$$

or $$4\pi \times 1{\cdot}6 \times 10^{-3} \text{ W} = 2{\cdot}02 \times 10^{-2} \text{ W per candle,}$$

or $$2{\cdot}02 \times 10^{5} \text{ erg per candle sec.}$$

So that $9{\cdot}73 \times 10^{7}$ erg will yield

$$9{\cdot}73 \times 10^{7}/2{\cdot}02 \times 10^{5} \text{ candle-sec ,}$$

or $$4{\cdot}82 \times 10^{2} \text{ candle-sec.,}$$

which is the desired datum. It may be regarded as a kind of theoretical maximum for integrated nitrogen after-glow.

We have now to consider what fraction of it is actually yielded in practical cases. It is scarcely necessary to say that in such cases the nitrogen will *not* be at N.T.P., but highly rarefied

There is a technical difficulty in getting the brightest possible after-glow, combined with the longest duration. For long duration we require to have the walls coated with sulphuric, or phosphoric acid, to minimize the wall effect (Rayleigh 1935, p. 569). But a vessel so coated is unsuitable for generating active nitrogen, because the discharge decomposes the coating, and contaminates the nitrogen. On the other hand, time is required to pass the excited gas from a clean to a coated vessel, and during this time the greatest brightness which contributes largely to the integrated total is lost. A compromise was adopted, which practically amounts to coating part of a composite vessel. It was found satisfactory. A small flask of 300 c.c. capacity with clean glass walls was used for generating the active gas. It was fastened with plasticine into the neck of a large globe 19 l. capacity (figure 1). The plasticine was applied outside, in the re-entrant angle, so that the gas had little access to it. The neck of the small flask, which projected into the large one, was about 4·3 cm. in diameter.

The glowing gas was generated in the smaller vessel, and passed into the larger one by diffusion, and by a kind of pumping action caused by the intermittent discharge. The walls of the larger vessel were coated with gummy phosphoric acid, and some undeliquesced phosphorus pentoxide lay on the bottom. The coated walls, kept dry in this way, are favourable. The uncoated walls of the smaller vessel are less favourable, but the area of this is comparatively small, and the aggregate result was satisfactory, the glow once generated remaining visible for several hours.

Preliminary photographic tests were made to determine what was the best pressure to use, in order to get the highest ratio of integrated light to gas density. The plate was exposed for the whole duration of the after-glow, and integrated the light automatically, well enough at any rate for approximate purposes. The number of excitation given was inversely as the pressure. In this way it was found that a pressure in the neighbourhood of 0·2 mm. was best with the discharge arrangements used, and this pressure was adopted.

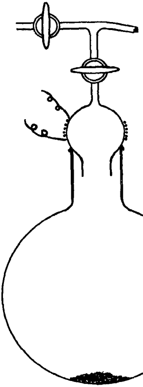


FIGURE 1

The light of the exciting discharge was cut off by covering the small bulb, and also the neck of the large globe with opaque material, so that only the after-glow light in the large bulb was able to emerge horizontally. The initial intensity could thus be determined while the exciting discharge was running, which is advantageous. Although photometric observations can readily be made as the glow decays through a given intensity, it would be difficult to begin observation at the exact moment the discharge was turned off; and the first few seconds are all-important.

For the photometry it was useless to attempt any great refinement, and the simple method of Bouguer's photometer with a diffusing screen was used. A wooden cube had two adjacent sides coated with white paper, the corner forming the photometric edge (figure 2). One side was illuminated by the after-glow bulb, and the other by the standard source, a small incandescent lamp measured at the National Physical Laboratory.

The following was a typical experiment:

Time (sec.)	Candle power
0·0	1·06
2·4	0·359
4·2	0·203
6·4	0·090
9·8	0·051
22·4	0·013

From these data the time integral was determined graphically as

3·18 candle-sec.

Although the luminosity in the bulb remains visible for hours, its equivalent duration at the initial intensity is only about 3 sec. The prolonged faint luminosity contributes very little to the total.

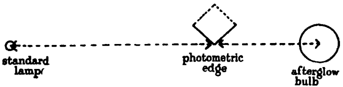


FIGURE 2

In the above experiment the gas pressure was 0·215 mm. of mercury, and the volume 19 l. Thus the volume of gas reduced to N.T.P. is 5·08 c.c. and the light emitted 0·626 candle-sec. per c.c.

We have seen that if every molecule present gave out a quantum, the light would amount to $4\cdot82 \times 10^{2}$ candle-sec. per c.c. The actual emission is $1\cdot3 \times 10^{-3}$ of this.

I do not think it would be easy to get a very much larger fraction. It is believed that the wall effect was small, and the initial glow was very bright, judging from previous general experience. The only direction in which there is possible scope for improvement is in the use of more powerful discharges. This would require costly equipment, and it is doubtful whether the brightness would be much increased. Even if, contrary to expectation, the light were increased a few times, it would still remain a small fraction of the calculated maximum.

2. INCREASING THE CONCENTRATION OF INERT NITROGEN, THE CONCENTRATION OF ACTIVE NITROGEN REMAINING CONSTANT

✓ The question of whether collisions with ordinary nitrogen molecules enter into the glow process suggests the experiment of adding nitrogen to the glowing gas, and observing whether any increase of luminosity results. I tried this experiment many years ago (Strutt 1912, p. 266), but observed no increased luminosity, unless the nitrogen was introduced in such a way as to cause local compression. It was subsequently repeated by Bonhoeffer and Kaminsky (1927) who also got a negative result. Later, the same experiment was carefully made by Kneser (1929) who was able to observe an increase of luminosity, and interpreted it as proving that the tempo of the whole process was in fact conditioned by collisions with neutral molecules.

I have not hitherto been completely satisfied by Kneser's results (Rayleigh 1935, p. 583), partly because of the earlier negative results mentioned, but mainly because the results of experiments on the compression of the glowing gas appeared inconsistent with it. I have recently returned to the question and am now quite satisfied that Kneser is right. ✓

It appeared that if such a test was made in a large vessel under conditions where the decay was so slow that it did not make much difference from minute to minute, the enhanced luminosity, if real, must come into evidence, and remain visible for so long that local compression of active gas in the vessel could not possibly explain it. It was thought best not to aim at spectacular effects, but to design the experiment so as to give results of quantitative value.

In this work the uranium photometer was used which has already been described (Rayleigh 1935, p. 574). It is enough to mention that the luminosity of the gas is matched against the (weak) luminosity of potassium uranyl sulphate. This latter is stimulated by its own radioactivity, and is constant in intensity. Equalization is effected by a series of calibrated neutral glasses.

Two series of experiments were made at different times. In one, a glass bulb of 12 l. capacity was used, which had been washed out with weak hydrofluoric acid, allowed to drain, and subsequently exhausted over phosphorus pentoxide. This bulb was uncoated The discharge was passed in it, at a low pressure, and the after-glow observed photometrically. When it had sunk to the same intensity as the potassium uranyl sulphate in the photometer more nitrogen was immediately admitted from a space between two taps, so as to increase the pressure about fivefold. The gas brightened up as described by Kneser, and the intensity was measured as soon as possible, and its further decay observed. The following is a sample experiment:

			↓			
time (min.)	0	2·4	6·5	7·1	9·2	13·4
$\log_{10}$ intensity	0·6	0·3	0·0	0·7	0·3	0·0
			↑			
			Gas admission			

The initial pressure was $3{\cdot}4 \times 10^{-3}$ cm , and the final pressure $17{\cdot}4 \times 10^{-3}$ cm.

Increase of nitrogen pressure	5·1 times.
Increase of light	5·0 times.

Later the experiments were repeated, using the 19 l. globe coated with phosphoric acid, with discharge annexe uncoated (figure 1). This gave a slower decay of the glow, and allowed more time for making the observations:

				↓						
time (min.)	0·0	8·6	16·4	26·0	28·0	28·8	29·7	31·0	32·1	33·2
$\log_{10}$ intensity	0·9	0·6	0·3	0·0	0·5	0·4	0·3	0·2	0·1	0·0
				↑						
				Gas admission						

Initial pressure $2{\cdot}7 \times 10^{-3}$ cm. Final pressure $12{\cdot}8 \times 10^{-3}$ cm.

In this case, no attempt was made at a very quick observation of the increased intensity, but a small extrapolation backwards in time was made to determine it (see figure 3).

It was found that

Increase of gas pressure was	4·73 times.
Increase of light	4·5 times.

These experiments were made at very low intensity, the brightness, viewed along a diameter of the bulb, being about 5×10^{-9} candles per sq. cm. which is near the minimum for satisfactory photometric observation. They indicate that the intensity is increased in the same ratio as the nitrogen pressure. Experiments made by admitting nitrogen to an after-glow some 9 or 10 times brighter than this gave only about a twofold increase of light when the nitrogen pressure was increased 5 times.

Kneser's results were similar. He obtained (Kneser 1929, p. 244) about a threefold increase of light when he increased the pressure tenfold, from 5×10^{-3} cm. to 5×10^{-2} cm. Evidently some disturbing cause comes in which, it may be suspected, is more conspicuous with higher gas pressures and glow intensities. Probably the same cause would account for the negative results of the earlier qualitative experiments on this effect: but I have not attempted to probe it further, because other problems seemed more urgent.

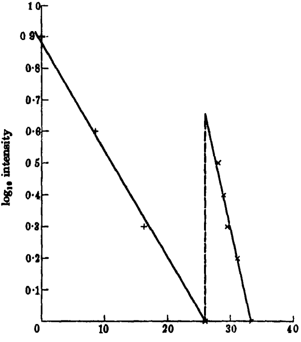


FIGURE 3

3. DIMINISHING THE CONCENTRATION OF ACTIVE NITROGEN, THE CONCENTRATION OF INERT NITROGEN REMAINING CONSTANT

In the experiment just described, we vary the concentration of inert nitrogen, maintaining the concentration of active nitrogen constant. I now come to experiments in which the converse charge is made, the concentration of inert nitrogen being maintained approximately constant, while the concentration of active gas is varied. This is done by allowing active nitrogen to diffuse from a smaller to a larger volume of inert nitrogen.

Two flasks (A, B, figure 4), of 1 litre volume each, were moistened with strong sulphuric acid, and fitted by means of ground joints to the two sides of a stopcock, C, of 1 cm. bore. The electrodeless discharge can be passed in the auxiliary vessel, D. In making an experiment, the whole is filled with nitrogen at low pressure. C is closed, E remaining open. The discharge is passed in D, and active nitrogen passes into A by diffusion. The discharge is stopped, and the temperature and pressure allowed to settle down to uniformity. E is closed and the glow in A is allowed to decay to a suitable intensity, as measured by the uranium photometer. The moment the chosen intensity is arrived at, C is opened. The glow then quickly diffuses into B, and under the conditions used, this occurs quickly enough for the decay during the time of diffusion to be of small importance. The intensity in A diminished, and in B it increases, equality being reached in a few seconds. The light in B is watched with the photometer, and its maximum intensity noted.

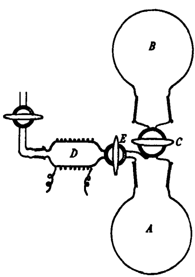


FIGURE 4. (⅛ actual size.)

In making a series of such tests, it would be possible after an experiment to wait for the glow to fade away completely, and then to excite again: but this is wasteful of time. It is better to exhaust and recharge with fresh gas

The measurements, which have to be made quickly, are somewhat difficult to carry out at the low intensity of the uranium photometer, and a photometer provided with a phosphorescent screen about 5 to 8 times as bright as this was used. It was of Balmain's paint, exposed to daylight some hours previously, and remained sensibly constant during a test, which is all that is strictly required. Moreover, it remained pretty constant during a set of tests, so that the conditions of each test did not vary greatly from that of the others of a set.

First series Nitrogen pressure 10^{-2} cm.

$\log_{10}$ initial intensity	0·6	0·7	0·6
$\log_{10}$ final intensity	< 0	> 0	< 0

Here we have $\log_{10}$ intensity ratio $=$ 0·65 say

Intensity ratio $=$ 4·5

Second series. Nitrogen pressure 7×10^{-3} cm.

$\log_{10}$ initial intensity	0·6	0·5	0·7	0·6
$\log_{10}$ final intensity	0	< 0	> 0	0

$\log_{10}$ intensity ratio $=$ 0·6

Intensity ratio $=$ 4·0

Third series. Nitrogen pressure 35×10^{-3} cm.

$\log_{10}$ initial intensity	0·5	0·6	0·7	0·6	0·6	0·6	0·7
$\log_{10}$ final intensity	<0	<0	>0	<0	<0	0	>0

$\log_{10}$ intensity ratio = 0·63 say
Intensity ratio = 4·3

In discussing these experiments it is assumed that the active gas, at these low luminosities, is a negligible fraction of the whole. This being so, we conclude that if the active gas contained in 1 litre of rarefied nitrogen is allowed to diffuse quickly into 2 litres the candle power per unit volume is diminished about 4·3-fold. This is in fair agreement with the value 4, which would be expected if the reaction were bimolecular as regards active nitrogen. It is thus in agreement with the conclusion which has generally been drawn from observations on the rate of decay of the luminosity in a closed vessel.

4. INCREASING SIMULTANEOUSLY THE CONCENTRATION OF ACTIVE AND INERT NITROGEN

The results arrived at so far allow us to predict what should be the effect of compressing the glowing gas. Let unit volume be suddenly compressed to volume $1/n$. This will increase the concentration of (1) inert nitrogen, and (2) the active nitrogen. The initial luminosity should be increased n times by (1) and n^2 times by (2). So that on the whole it should be increased n^3 times.

This is not in agreement with my own experiments, which appear to be the only ones hitherto made (Rayleigh 1935). These experiments were more consistent with an n^2 law, and the brightness certainly did not increase as much as an n^3 law would require.

I have returned to the question in the hope of clearing it up.

In the former experiments the piston which compressed the gas was a rising column of mercury, covered with a layer of sulphuric acid. It was found that sulphuric acid alone would not do, for the glow was extinguished as soon as the acid was caused to rise. This was attributed to gaseous impurities coming out of the moving acid. It is possible that even a film of acid may have acted in the same direction. Moreover, a liquid piston is open to the objection that it cannot well be started very quickly from rest; thus the compression takes time, and the question of decay of luminosity enters. It is true that in the previous work this was judged not to be of much practical importance: but in returning to the problem I decided to use a solid piston, which can be quickly moved from its initial to its final position in the cylinder.

Figure 5 shows the arrangement. *A* is the glass compression cylinder. Glowing gas diffuses in at *B* from a store in a large bulb similar to that shown in figure 1. The piston *C* is of indiarubber coated with gummy semi-deliquesced phosphoric acid, which acts as a lubricant, keeping it tight, and helped it to move freely in the cylinder. The piston rod of glass moves up through a barometric column as shown. The handle *D* controls it. When enough glowing gas has diffused in, the piston is raised to a standard position just above the post *B*, isolating the gas in the cylinder. At this stage the handle *D* rests on the lower of the two stops (not shown). Compression was always to half volume, and it was made by raising it to an upper stop. This operation could be carried out by the assistant in a fraction of a second. The stopper *E* at the top is for the introduction of the metaphosphoric acid lubricant. The same material serves to cement the stopper airtight. The whole cylinder is coated internally with metaphosphoric acid, to minimize the unfavourable effect of the walls on the glowing gas The photometer is applied near the top, its field being shown by the dotted square.

In working with this apparatus, it is necessary to make sure that the piston has enough lubricant on it for tightness The test of this is to raise it above the point *B*, and observe whether the glowing gas is able to diffuse past it. If not, the piston may be considered tight. This is quite a sensitive test, and is directed to the exact point at issue.

The best experiments with this apparatus were made with an initial intensity about 4 times that of potassium uranyl sulphate. Since, however, we are looking at only 3 cm. thickness of the luminous gas, this represents a much greater intrinsic luminosity of the gas than was used in looking along the diameter of large flasks or globes in some of the previous experiments. The initial nitrogen pressure was $2{\cdot}7 \times 10^{-3}$ cm and it was compressed to twice the amount

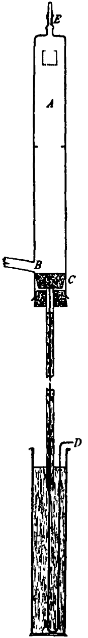


FIGURE 5.
(⅛ actual size.)

$\log_{10}$ initial intensity	0	0	0	0	0	0	0	0	0
$\log_{10}$ final intensity	$>0{\cdot}8$	0·9	0·9	$>0{\cdot}8$	$<1{\cdot}0$	$>0{\cdot}8$	0·9	$<1{\cdot}0$	$<1{\cdot}0$

$\log_{10}$ intensity ratio = 0·9
Intensity ratio = 7·9

Thus the experiments confirm, as accurately as could be expected, that the brightness varies as the inverse cube of the volume. As already shown, this result is the logical conclusion from, and confirmation of, the previous experiments in which the concentration of (1) the inert nitrogen and (2) the active nitrogen are separately varied.

5. DIMINISHING SIMULTANEOUSLY THE CONCENTRATION OF ACTIVE AND INERT NITROGEN

Notwithstanding this apparently satisfactory result, it was thought desirable to try to get a further experiment linked with the others. That is, expansion of the glowing gas as a whole, instead of compression. The method used for carrying this out does not differ much from that used for the diffusion experiment. The apparatus is the same (figure 5). The only difference is that the flask *B* is exhausted beforehand with cooled charcoal, so that on establishing communication between *A* and *B* the glowing gas passes out into a vacuum instead of diffusing into a nitrogen atmosphere. The photometer was used with the low luminosity of potassium-uranyl sulphate.

Initial pressure 2×10^{-2} cm.

$\log_{10}$ initial intensity	0·9	0·9	0·8	0·9	0·9	0·9	0·9
$\log_{10}$ final intensity	0·1	0·1	0·0	0·1	0·1	0·1	0·1

$\log_{10}$ intensity ratio = 0·8
Intensity ratio = 6·3

Initial pressure 5×10^{-3} cm.

$\log_{10}$ initial intensity	0·9	0·9	0·9	0·9
$\log_{10}$ final intensity	0·05	0·05	0·05	0·10

$\log_{10}$ intensity ratio = 0·85
Intensity ratio = 7·1

The diminution of brightness is definitely less than was expected. The intensity ratio should be 8·0 and it is not found in any single experiment to be more than 7·1. In other words, after the expansion the gas is *brighter than it ought to be in* the ratio of at least 1·12. That it really was too bright was checked by putting the eightfold reduction on the photometer, and observing that in fact the expanded gas was brighter than the standard light so reduced. It is difficult to explain this result or to see how it can be fitted into the scheme of thought which serves to co-ordinate most of the others. Most of the complications which suggest themselves as possible would tend to make the gas less bright than the simple theory would lead one to expect. But here the final brightness is definitely too great, and it seems certain that there is a residual effect unexplained.

6. Effect of temperature change

It was observed at an early stage of my investigations on active nitrogen that the glow became brighter at low temperatures (Strutt 1911, p. 221). Further, it was established that this effect could not be wholly explained as the result of increased concentration of the active material in the cooled part of the vessel (Strutt 1911, p. 262). It was, as pointed out at the time, a unique example of a reaction velocity being accelerated by lowering of temperature (Strutt 1911, p. 264).

Since then, a few other examples of such reactions have come to light. In particular, Bodenstein's (1922) measurements of the rate of oxidation of nitric oxide by oxygen have shown a negative temperature coefficient. It is of interest that this is one of the few examples of a termolecular reaction, and Bodenstein interprets his result as due to an increase in the time during which two of the partners remain in association which increases the chance of collision with the third partner.

Since the evidence collected in the present investigation points strongly to a termolecular reaction in the reversion of active nitrogen to ordinary nitrogen, I was led to a quantitative measurement of the temperature effect, such as could hardly have been carried out with the technique of the earlier period.

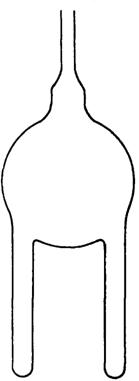


FIGURE 6.
(½ actual size.)

A vessel of 300 c.c. capacity, shown in figure 6, was connected to the 19 l. globe. It had two legs 10 cm. long and 1·5 cm. internal diameter. The glow passed into this vessel by diffusion. One of the legs was kept at the ordinary temperature (15° C) while the other was heated or cooled. To take the most extreme case when the cooling was in liquid oxygen (– 183° C) it was found that the brightness which was greatly enhanced by the low temperature was uniform throughout the distance of 7 cm. from the bottom, of which 5 cm. was immersed. It may be inferred from this that even the lowest part of the leg was still receiving a free supply of active nitrogen from the bulb above, and that the pressure equilibrium of this substance was not appreciably disturbed by its removal by decay in the upper part of the leg before it could reach the lower part. It is judged therefore that active nitrogen as well as the ordinary nitrogen with which it is mixed, is in pressure equilibrium in the bulb and the two legs.

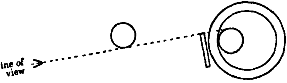


FIGURE 7

The measurement consisted in comparing the intensity in the leg at room temperature with that in the other leg which was maintained at a higher or a lower temperature. The method of comparison was to view the legs in a direction which would bring them into approximate juxtaposition (figure 7). They were then equalized by making use of a series of neutral tinted glasses of graduated and measured density. A glass was held in front of the brighter limb. The density ($\log_{10}$ opacity) of these glasses increased by steps of 0·1, and the measurement was made to 0·05, representing about 12 %.

Considering the large difference of brightness, this simple photometric method was adequate The temperatures used were 100° C (a bath of water just boiling), – 78·2° (solid carbon dioxide in alcohol) and – 183° C (liquid air). The results were as follows:

Temperature (° Abs.)	Luminosity	Luminosity reduced to uniform concentration
373	0·447	0·97
288	1·00	1·00
195	4·47	1·39
90	79·4	2·42

For experimental reasons, the brightness has been compared at a given pressure, rather than a given concentration. Before we can infer anything about the reaction velocity at different temperatures it is necessary to correct the results to what they would be if the local concentration in the cooled leg of the vessel did not occur. The concentration is taken to be inversely as the absolute temperature, and the light to vary as the cube of the concentration, in accordance with the general results of the earlier sections of this paper. The intensity at a temperature T will therefore be reduced to standard by applying a factor $(T/288)^3$. This is done in the last column, and it appears that the very striking difference of brightness which follows on cooling is largely though not wholly due to the increased local concentration. There is a definite negative temperature coefficient, even when the increased concentration has been allowed for. The same thing was shown qualitatively many years ago by wholly immersing a closed bulb containing glowing nitrogen in liquid air, local changes of concentration being thus avoided The glow was brighter in this cooled bulb than in a similar bulb initially excited to the same brightness, which was not cooled, though the light did not last so long.

The increase of reaction velocity with diminishing temperature which remains when the correction has been made is not very great. and in this respect the results resemble those of Bodenstein on the oxidation of nitric oxide. Over the extreme range investigated, his rates of reaction vary as $T^{-1\cdot 53}$. The corresponding power for the present results is $T^{-0\cdot 64}$.

The positive temperature effect on the rate of ordinary chemical reactions is enormously more marked than the negative effect on either of these two termolecular reactions. It often corresponds to a power T^{+100}, or even more.*

I have much pleasure in thanking my assistant, Mr R. Thompson, for the help he has given me in this work

REFERENCES

Bodenstein, M. 1922 *Z Phys. Chem* **100**, 68.
Bonhoeffer, K. F. and Kaminsky, G. 1927 *Z. Phys. Chem.* **127**, 385–400
Ives, H. E. 1924 *J. Opt Soc. Amer.* **9**, 635–638.
Kichlu, P. K. and Acharya, D P 1929. *Proc Roy Soc* A, **123**, 168–171.
Kneser, H. O. 1929 *Ergebn. Exact Naturw.* **8**, 229-257.
Rayleigh 1935 *Proc. Roy Soc.* A, **151**, 567–584.
Strutt, R. J. (Rayleigh) 1911 *Proc. Roy. Soc* A, **85**, 219–229.
— 1912 *Proc. Roy. Soc.* A, **86**, 262–269
Walsh, J. W. T. 1926 *Photometry*. London

* In studying this subject I have derived much assistance from Hinshelwood's *Kinetics of chemical changes in gaseous systems* (3rd ed., Oxford, 1933).

New studies on active nitrogen

II. Incandescence of metals in active nitrogen, and quantitative estimates of the energy liberated

By Lord Rayleigh, F.R.S.

(*Received* 3 *May* 1940)

It is shown that pieces of sheet gold, copper, silver, or platinum may be made red hot or even melted by exposing them to active nitrogen produced in a low-pressure discharge. The nitrogen gives up its energy to the metal, which remains unacted on.

This phenomenon is applied to measuring the energy of active nitrogen, drawing a known quantity of gas through the activating discharge and then over the metal, and measuring the energy liberated on the latter.

The amount of energy collected from the gas was surprisingly large, and is difficult to reconcile with existing theories of the nature of active nitrogen. In some cases the energy was as high as 10 eV for every molecule of nitrogen that passed through the discharge. This quantity of energy can with difficulty be accounted for by dissociation, even if it occurred to the extent of 100 %. The energy radiated as after-glow under favourable conditions is only of the order of 10^{-8} of the energy collected by the metal.

Experiments will now be described which show that a surprising amount of energy can be collected from nitrogen which has been made active in a low-pressure electrodeless discharge. This quantity of energy is very large compared with the energy of the after-glow light and can only be accounted for with difficulty if it is attributed to dissociation with the accepted value of 7·34 V energy. It becomes necessary to make the extreme supposition of 100 % dissociation, and even this hardly suffices.

I have observed that pieces of metal sheet hung up in a large discharge vessel in which nitrogen was made active by the electrodeless discharge at, say, 0·3 mm. pressure, will become red hot, and even melt and run down. The arrangement shown in figure 1 is convenient for exhibiting this effect. A large bottle is provided, with a side entry for exhaustion and admission of gas, and a stopper from which the specimen can hang, using preferably a hard glass or silica hook. At the bottom of the bottle is a coil for exciting the electrodeless discharge, and the specimen hangs high above it, so as to be far from the ring discharge. If any difficulty is experienced in getting the action to start, the coil may be raised so as to bring the discharge closer for a moment, but this is usually unnecessary. The gas is made active and diffuses to the metal, when it gives up its energy in reverting to the normal condition. A piece of sheet gold or silver 1 cm. square may readily be got hot enough to melt and run down. Under suitable conditions it can be observed that the luminosity in the gas is diminished in the neighbourhood of the hot metal.

The metals gold, silver and platinum will all do this without tarnish of the surface. Gold and silver can readily be melted. They appear to be somewhat more active than platinum. Copper, if initially oxidized, becomes clean and has on occasion been melted. The surface of a compact rolled sheet of copper after it has been made red hot in this way, becomes red and spongy, like an electrolytic deposit, but can be made compact and lustrous again by burnishing.

Comparisons of different metals have been made by hanging two strips side by side, so that they were exposed in a manner strictly comparable Often one of the metals would become incandescent while the other remained dark. This was due to "hang fire", and nothing could be inferred until both strips were visibly red hot. Attending to this condition, the result of a series of comparisons was to place the metals in the following order of brightness: gold, copper, silver, platinum. Under comparable conditions the earlier in the list always becomes brighter than the later.

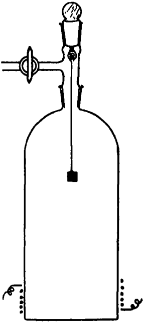


FIGURE 1.
(½ actual size.)

Iron and nickel have both been observed to become incandescent in active nitrogen. These metals become covered with a black coating. Most attention has been given to nickel. This metal easily gets red hot, but becomes duller after a time, presumably owing to the protective action of the coating, and cannot then be started again after cooling.

There is sometimes difficulty in getting the action to start. The main secret is to have the surface quite clean. Gold and platinum are best cleaned by chromic acid and distilled water, followed by ignition in a spirit flame. After a specimen of noble metal has glowed once, it will usually do so again on restarting the discharge after an interval. If the specimen is hung up by a clean platinum wire, say 0·15 mm. diameter, the wire starts glowing and the action is communicated to the metal foil hung from it, no doubt by the local heating. Or the platinum may be heated electrically to start the action, the heating current being switched off afterwards. These devices are useful when the supply of active nitrogen is restricted. When there is free diffusion from a discharge in the same vessel they are scarcely necessary.

✓The chemical action on metals like iron and nickel has not yet been successfully examined. Presumably the black coating consists of nitride, as in the cases of zinc, mercury and other metals (Strutt 1913, p. 542). In the present case, the volume of nitrogen absorbed is small, and the ammonia reaction has not been obtained from it. But the formation of a black coating is definite and indeed striking ✓

✓In case of metals which remain bright, the action may clearly be regarded as catalytic, the active nitrogen being converted and giving up its energy by a process which occurs at the metal surface. The effect can be taken advantage of to investigate the energy liberated.* ✓

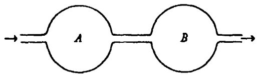


FIGURE 2

Gold was selected as the most active metal, or at any rate the most active of those which remain bright.

In the original form of experiment, in which the metal is hung up in the discharge vessel though away from the discharge, it is not practicable to determine how much of the gas is transformed per second. The same gas is repeatedly activated and deactivated as it diffuses from the discharge and back again. For quantitative work it is necessary to have a determinate stream of gas exposed to the action of the gold, which cannot get back to the activating discharge. This is not so easily realized as was at first supposed, and many early experiments had to be discarded for this reason. A thorough investigation of the point was made before the final experiment was designed.

Suppose that a stream of nitrogen passes at low pressure from discharge vessel A to an after-glow vessel B through a connecting tube (figure 2); we have to consider what conditions the length and diameter of the tube must satisfy to prevent the gas in B after deactivation from diffusing back to A and getting activated a second time This question may be attacked theoretically or experimentally

---

* Willey (1927) has observed that copper and other metals become slightly warm in a stream of active nitrogen, and has applied this to quantitative energy measurements; but the energy collected from a given quantity of nitrogen in his experiments was comparatively very small; of the order of a thousandth part of that here obtained His experiments were made at a much higher gas pressure

Beginning with the theoretical discussion we make use of Einstein's result (Einstein 1905, p. 559) that the distance of mean square that a gaseous particle is carried by diffusion in time t is $\sqrt{(2Dt)}$, where D is the coefficient of diffusion We are concerned, however, not with the distance of mean square, but with the greatest distance reached by any appreciable percentage of particles. This is of course indefinite, but not more so than the question itself, as formulated above. If this distance is $n\sqrt{(2Dt)}$ we may take $n = 2$, which should leave only 5 % of all particles outstanding, or $n = 2{\cdot}4$, which should leave less than 1 % outstanding. The few that get farther are of no practical importance for our purpose.

If diffusion is occurring in the face of a gas stream of velocity u, the net distance travelled in time t is

$$n\sqrt{(2Dt)} - ut,$$

and this will be a maximum if

$$\frac{d}{dt}(n\sqrt{(2Dt)} - ut) = 0,$$

which gives

$$t = n^2D/2u^2,$$

and the greatest distance, l, which a particle can penetrate upstream is given by

$$l = n^2D/2u = n^2D_0 \times 760/p$$

if D_0 is the coefficient of diffusion of active nitrogen into ordinary nitrogen, or into itself, at atmospheric pressure, and p is the (low) pressure in mm. of mercury.

If v is the volume of gas taken in per second at atmospheric pressure, then the expanded volume is

$$v \times 760/p,$$

so that if d is the diameter of the tube

$$\tfrac{1}{4}\pi d^2u = v \times 760/p,$$

consequently the necessary length of tube is given by

$$l > \pi d^2 n^2 D_0 / 8v,$$

or

$$v > \alpha d^2 / l,$$

where

$$\alpha = \pi n^2 D_0 / 8.$$

We do not know the value of D_0, but from comparison of known coefficients in other cases, a value of 0·2 may be assumed. If $D_0 = 0{\cdot}2$, and $n = 2{\cdot}4$, $\alpha = 0{\cdot}45$ and

$$v > 0{\cdot}45 d^2 / l.$$

To investigate the question experimentally the general arrangement of figure 2 was used, but the downstream vessel *B* was used for the discharge, and the vessel *A* was used to observe whether or not the glowing gas had been able to diffuse back against the stream. If a glow appeared in *A*, then it followed either that the length of the connecting tube was too short, or that the linear velocity of the gas was too small.

The vessels *A* and *B* were 5 l flasks, and in different experiments tubes of 1, 1·5, 2 and 3 cm bore, and from 10 to 50 cm. in length, were used to connect them. At each set up, the rate of intake was varied by means of the porous tube valve to be presently described. Below a certain rate of intake the glow was perceptible in the vessel *A*, and at lower intakes still it became conspicuous. Nothing would be gained by describing these tests in further detail. They are summed up by saying that within the limits of experimental error the condition for no glow in *A* was found to be

$$v > d^2 / l.$$

That the numerical coefficient of the expression on the right is unity is of course a coincidence

Thus there is agreement between the theoretical and experimental criterion, as regards the order of magnitude, which is all that can be aimed at: for we have no definite measurement of the coefficient of diffusion used in the calculation, or of the minimum amount of back diffusion observable in the experiment. The criterion determined by experiment, which is the more severe, requiring a tube about twice as long, or a larger admission of gas, has always been adopted in practice.

The nitrogen used was taken from a commercial cylinder with a pressure regulator (Beard's pattern) and thence through a red hot silica tube 35 cm. long and 4 cm. diameter filled with iron wire of No. 30 s.w.g. The gas then passed over phosphorus pentoxide, and in some cases it passed through a trap cooled with liquid air: but this latter made no observable difference, and was usually omitted. To regulate the stream, on passage from atmospheric pressure to low pressure, it passed by diffusion through the walls of a porous tube (clay tobacco pipe stem) which could be shortened at pleasure by partially immersing it in mercury (Wansbrough Jones 1930, p. 533). The porous tube was graduated in centimetres, and the rate of intake at each mark was determined by preliminary calibration, taking the gas in from a graduated vessel. This porous tube arrangement was found more dependable than various forms of needle valve that were tried.

In the definitive experiments, the discharge vessel was sometimes a bulb of 250 c.c. capacity, sometimes a cylinder of glass 8 cm. long and 4·7 cm. internal diameter (figure 3). In either case the exciting coil consisted of six turns of wire. An air blast kept down the heat of the discharge, which was considerable. The observation vessel was a glass bulb of 500 c.c. capacity, and the gold specimen was supported in the middle of it by a thin piece of platinum wire. If necessary the platinum wire could be heated by a current to start the action. The current was then discontinued

The gold could readily be withdrawn for cleaning or alteration. It sometimes became amalgamated during the intervals when it was standing idle, taking mercury vapour from the pump or the gauge. This was found to be a sign that it was in good condition, but the action seemed rather capricious. If amalgamation had occurred, the mercury was quickly driven off by the heat generated, when the experiment was started.

The connexion between the discharge vessel and the observation vessel was by standard ground joints, which made it easy to vary the length of connexion by putting in a supplementary piece.

A side opening from the observation vessel was connected to a McLeod gauge, and the bottom opening was connected through a tube containing copper gauze to a two-stage mercury vapour pump backed by a mechanical pump. The copper gauze destroyed any remaining active nitrogen and protected the mercury pump from being fouled by it. The pressure was from 0·3 to 0·6 mm. in different experiments, and the opening between the two vessels was wide enough to make the pressure in the discharge vessel nearly the same even when the flow was in progress.

The temperature of the gold was indicated by readings with an optical pyrometer. As will be explained, nothing depends on the absolute value of these temperatures.

On starting the discharge there was usually some "hang fire" or delay before the gold came into action, and a bright after-glow was seen in the bulb round the gold, and in the tube leading to the pump. As soon as the gold warmed up to its work, the glow was diminished to a fraction of its original intensity.

To interpret the results quantitatively we require the energy which is dissipated by a square centimetre of gold at the brightness observed by the optical pyrometer in the nitrogen work, and this was determined in each case by an independent experiment, in which a strip of gold was heated electrically. The arrangement is shown in figure 4.

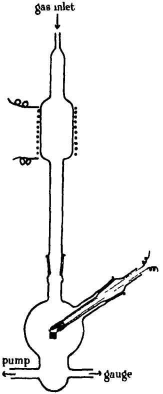


Figure 3

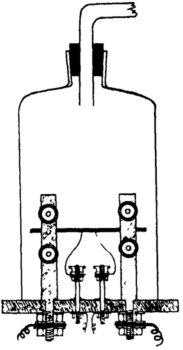


Figure 4

The strip cut from a gold sheet 0·1 mm. thick, had two gold wires of 0·08 mm diameter welded to it near the middle. They came away at right angles to the strip, both from the same edge, and this distance apart was measured as 0·673 cm. The breadth of the strip in this region was 0·09 cm. The wires served as potential electrodes. The strip itself was held between clamps 5·5 cm. apart, through which the current was led in. The clamps and connexions passed out through an ebonite base plate. A bell jar covered them, all being made airtight with plasticene. The pressure was reduced to that used in the nitrogen experiment which it was desired to imitate, so as to have the same convective loss. It was found, however, that within the range of 0·3–0·6 mm. used, the variation was insensible, so that provided that a pressure was used within this range, it was useless to refine further.

The heating current was taken from a source of low voltage (4 V) which diminished the risk of accidental burning out.

The potential drop was measured on a unipivot millivoltmeter of high resistance. The area contained between the potential electrodes counting both sides of the strip is $2 \times 0{\cdot}099 \times 0{\cdot}673$ sq cm. or 0·1332 sq. cm.

Example. With a gas pressure of 0·4 mm. the reading of 780° C on the pyrometer was attained with 11·2 amp. and 64 mV dropped This makes the energy emitted 0·717 W Another test at 0·6 mm. gave 0·712 W Mean 0·715 W; or per square centimetre of gold 4·75 W.

The temperatures given are not the actual temperatures of the strip, which is not of course a full radiator. They are "brightness temperatures". The actual temperatures are of no importance for our purpose. The pyrometer is merely used as an indicator to show that the gold strip used in the energy measurement is at the same temperature as was the gold specimen when caused to glow by the nitrogen reaction. Even if the scale of the pyrometer were perfectly arbitrary this would be attained, assuming merely that the instrument behaved with consistency.

Before passing to the quantitative results, we note that

$$1\,\text{eV} = 1{\cdot}591 \times 10^{-12}\ \text{erg.}$$

Number of molecules in 1 c c. of gas at N T P $= 2{\cdot}75 \times 10^{19}$ Thus, to excite all the molecules in 1 c.c to 1 V requires $4{\cdot}375 \times 10^{7}$ ergs.

The following is a specimen experiment·

Gold foil measures 1×1 cm. Area (both sides included) = 2 sq. cm.

Length of channel 19 cm. Diameter 2 cm.

Rate of gas admission 0·211 c.c. per second, measured at atmospheric pressure.

Pressure 0·44 mm. (this datum is not strictly necessary).

Reading of radiation pyrometer 780° C.

Energy dissipated in radiation and convection in auxiliary experiment = 4·75 W per sq cm.

Thus energy dissipated by the actual area = 9 5 W.

Or per c.c. of gas admitted (measured at atmospheric pressure) the energy dissipated in radiation and convection is $\dfrac{9{\cdot}5 \times 10^{7}}{0{\cdot}211} = 45{\cdot}0 \times 10^{7}$ ergs.

Thus the observed excitation is $\dfrac{45{\cdot}0}{4{\cdot}375}$ V or 10·3 V.

With this example in detail, I now give a tabular view of all the experiments made.

Channel length cm.	Pressure mm.	Intake c.c. per sec.	Pyrometer reading °C	Watts per sq. cm.	Area of gold sq cm	Watts developed on gold	Watts per c.c. of gas per sec.	Excitation volts (at least)
27	0·37	0·178	780	4·75	0·84	3·99	22·4	5·12
27	0·30	0·148	740	4·58	0·84	3·85	26·0	5·94
27	0·44	0·211	780	4·75	0·84	3·99	18·9	4·32
19	0·44	0·211	900	6·80	0·84	5·71	27·1	6·20
19	0·44	0·211	780	4·75	2·0	9·50	45·0	10·3
19	0·51	0·240	740	4·58	2·0	9·16	38·2	8·73
22	0·36	0·148	730	4·17	2·0	8·34	56·4	12·9
22	0·42	0·178	730	4·17	2·0	8·34	46·9	10·7
22	0·49	0·211	730	4·17	2·0	8·34	39·5	9·03
22	0·56	0·240	720	4·11	2·0	8·22	34·3	7·84

In the above table the first column gives the length of the channel, 2 cm. diameter, separating the discharge vessel and the vessel in which the gold is hung.

The intake, given in column 3, is measured in c.c. per second at atmospheric pressure.

The fifth column gives the dissipation of energy per square centimetre, corresponding to the pyrometer reading, as determined by the auxiliary experiment. The sixth column gives the total area of the gold, both sides included.

The last column shows what the average excitation of each molecule must be *at the least* to account for the energy extracted from the gas in that particular experiment. These values are not uniform. For that, it would be necessary that every molecule should be fully excited, and should give up its energy to the gold, and the nature of the arrangement does not lead us to expect that this would be realized. Perhaps it may be suggested that only the highest values were significant, or worth recording. I have thought it best to give all the satisfactory experiments which have been made, so as to allow a broader judgement of the general tenor of the investigation.

It is found that the temperature attained by the gold is in general less the longer the connecting tube. If the connecting tube is very short, this may be explained by back diffusion and repeated activation of the gas, as in the case of stationary gas in a closed vessel. In the above experiments the criterion which has been laid down for the prevention of back diffusion has always been satisfied, though not always by a large margin, and it is believed that this effect does not enter to any appreciable extent. In a long channel the gas loses energy before it reaches the gold when the energy remaining in it is largely (but not completely) given up.

It will be observed that the energy collected is in all cases of the same order of magnitude as that calculated on the basis of complete dissociation, and in some cases exceeds it.

Even if the gas is completely dissociated when it leaves the discharge, some recombination must almost certainly occur on the surface and in the volume of the connecting tube apart from what occurs on the gold. And, moreover, the gold certainly does not collect the whole of what remains, for though the gas glow is very much reduced when the gold gets hot and comes into action, yet the gas is still perceptibly glowing when it leaves the observation vessel. It appears therefore very difficult to admit that the energy can be accounted for as dissociation energy, at any rate if the value of 7·34 V is adhered to (Hertzberg and Sponer 1934).

The gold in the experiment first quoted in detail is emitting 9·5 W of energy. This, at the point of maximum luminosity of the spectrum (and the after-glow spectrum has its intensity concentrated near this point) would yield 475 candle power, the luminous after-glow in the gas which would be obtained in a large vessel in the absence of the gold is only a fraction of a candle, and is therefore of quite a different order of magnitude

Nevertheless, the two things are intimately connected. The luminosity of the active nitrogen falls to a marked extent when the gold comes into operation. As already explained this is seen when the action of the gold "hangs fire". Alternatively the gold may be removed from a large observation vessel by a winch mechanism worked from outside, and the after-glow light in the vessel then becomes much brighter, and the increase may be determined. There are, however, technical difficulties about carrying this out satisfactorily. For quantitative comparison it is probably better to rely on the results of the first paper of this series.

It was then found that with the best excitation attainable, the integrated light emitted by 1 c.c. (measured at atmospheric pressure) was 3·18 candle sec., or, using the mechanical equivalent of light,

$$3{\cdot}18 \times 2{\cdot}02 \times 10^5 \text{ ergs} \quad \text{or} \quad 6{\cdot}45 \times 10^5 \text{ ergs.}$$

Comparing this with the $4{\cdot}50 \times 10^8$ ergs collected by means of gold we see that the after-glow light represents not more than $1{\cdot}4 \times 10^{-3}$ of the whole amount of energy present, a quite insignificant fraction

The recognition of this fact must lead us to regard the whole complex of phenomena in rather a different light from that which has been usual hitherto. It was the visual phenomena that led to the recognition that nitrogen in a special state was present, and that appreciable percentage concentrations were chemically active. It now appears that the luminous phenomena are very subordinate from an energy point of view, and that only a small percentage of all the active material can be concerned in them. In other words, they are now regarded as a by-product of the main processes.

We must now recognize that under no circumstances actually realized is the luminous emission an essential part of the main change in progress. It is still possible and indeed probable that it indirectly measures the rate of transformation in progress in the volume of the gas. This is the implicit assumption which has generally been made by previous writers, including myself, and it is the working hypothesis of the first paper of this series. However, now that it is established that the luminous emission accounts for very little of the energy transformation, a certain reserve becomes necessary

We have no readily available index of what is going on in the volume of the gas except the luminous phenomena, and since these account for very little of the energy, what becomes of the rest is a matter in which there must at present be some element of conjecture

We cannot make any progress without assuming qualitatively that *bright after-glow and high-energy content go together*, and there is ample reason for assuming so much. It is only when there is bright after-glow that we can visibly extract energy by introducing gold and making it red hot thereby: the brightest glows are only obtained immediately after the excitation by powerful discharges, and they progressively diminish with time, obviously because the energy in the gas is dissipated and, most conclusive of all, the gaseous glow is extinguished when the gold comes into action and removes the energy

The question which presses for answer is what becomes of the bulk of the energy, since it is not accounted for by the after-glow. Is it dissipated in the volume or at the surface of the vessel, acting like a gold surface, but less efficiently?

This certainly does happen to some extent, and it is only in vessels with specially treated walls that the gas in its special state (always recognized by its capacity to emit after-glow) will keep for an hour or more (Rayleigh 1935).

But what would happen if we were able to remove the nitrogen from the action of the walls altogether, as, for example, by increasing the size of the walls indefinitely, or if we imagined that the action occurred under the kind of condition found in the gaseous nebulae or in the upper atmosphere? Would the integrated amount of light be increased 500 times? I cannot think so. Experiments made to compare the law of decay of bulbs coated (*a*) with phosphoric acid and (*b*) with apiezon oil, showed that although the glow lasted very much longer in the former, yet the decay *in the early stages* was nearly the same in both, and it is in these early stages that the greater part of the emission of light occurs, and therefore, according to our assumption, most of the energy is dissipated I consider it practically certain therefore that there must be large dissipation occurring in the gas space without emission in the visible or photographic infra-red regions of the spectrum.

In conclusion, I wish to thank my assistant, Mr R Thompson, for efficient help with the experiments, and Professor S. Chapman, F.R S , who has kindly checked the theoretical reasoning.

REFERENCES

Einstein, A 1905 *Ann Phys , Lpz* , **17**, 549–560
Hertzberg and Sponer 1934 *Z phys Chem* **26**, Abt. B, 1–7
Rayleigh 1935 *Proc. Roy. Soc* A, **151**, 567–584
Strutt, R. J. (Rayleigh) 1913 *Proc Roy. Soc* A, **88**, 539–549.
Wansbrough-Jones, O. A 1930 *Proc. Roy. Soc.* A, **127**, 530 539.
Willey, E. J. B. 1927 *J. Chem Soc* **292**, 2188–2196

X-ray electrons expelled from metals by silver $K\alpha_1$ radiations

By C. J. Birkett Clews, Ph.D., and H. R. Robinson, F.R.S.
Queen Mary College (University of London)

(*Received* 3 *June* 1940)

The energies of the X-ray electrons expelled from targets of gold, platinum, silver and copper by silver $K\alpha_1$ rays have been measured in the magnetic spectrograph. The results have been combined with those of earlier measurements, and they yield a value for Planck's constant of $6{\cdot}62_4 \times 10^{-27}$ erg sec

The work to be described is a continuation of that published by us about five years ago (Robinson and Clews 1935). It consists of new measurements with the magnetic spectrograph constructed some time ago at Queen Mary College, and of their application to the problem of the atomic constants. Use has already been made of the fact that measurements of the photoelectron energies by the magnetic deflexion method, combined with a knowledge of X-ray characteristic frequencies, can yield a value of $e/m \times e/h \times \rho$, where ρ is the value of the X unit in 10^{-11} cm. ρ and e/m are now, it is thought, known with considerable accuracy, and our work may therefore be regarded as a determination of h/e There is still considerable disagreement among the results of different measurements of this quantity, and it seemed therefore desirable to extend as far as possible the range of our measurements with the magnetic spectrograph.

As our spectrograph was originally designed (Robinson, Andrews and Irons 1933) for work with relatively soft X-rays, the fields required in the present experiments with silver K-rays are appreciably outside the comfortable working range of the instrument. In fact, the difficulties mentioned in paragraph 3 of our earlier paper (Robinson and Clews 1935) were all encountered in an even more acute form. We were, however, able, by running for relatively short periods, with moderately long intervals for cooling the Helmholtz coils, to obtain satisfactory photographs It is obviously not practicable to work with appreciably harder primary radiations, fortunately, the general nature of the results now obtained indicates that there is no need to do so.

The new measurements are on the X-ray electrons expelled by silver $K\alpha_1$ rays from the M levels of gold (79) and platinum (78), the L levels of silver (47) and the K level of copper (29). For all but the last of these we have available for comparison results with chromium $K\alpha_1$ rays (Robinson 1934*a*), copper $K\alpha_1$ (Robinson *et al.* 1933; Robinson 1934*b*) and molybdenum $K\alpha_1$ (Robinson and Clews 1935). The copper K electrons can only be compared with those ejected by the molybdenum $K\alpha_1$ rays, the other radiations used being too soft to eject these electrons. In the comparisons the "difference" method introduced by Robinson (1934*b*, p. 1097) is used throughout. In this method the only X-ray data required are the values of the primary $K\alpha_1$ frequencies, which are very accurately known in terms of Siegbahn units. No use is made of X-ray spectroscopic values of the critical absorption frequencies, which are much less accurately known. Further, the method completely eliminates errors due to surface work and reduces the effect of other systematic errors

It should be emphasized here that all the measurements used in the following comparisons were made with the same magnetic spectrograph, and that the same plate-holder, of rigid construction, was used throughout. The quantity measured in the experiments is the product (rH), where r is the radius of curvature of the path of an X-ray electron as it moves in a magnetic field of strength H normal to this path. In the earlier work the magnitude of H was so adjusted as to bring r very close to a standard value (5 cm) for each line measured In the present work it was just not possible to do this, the required field being, except for the copper K electrons, just outside our attainable limit. The correction to standard radius is, however, small, and can easily be computed The accuracy of the correction was confirmed by additional checks, using the very simple device of employing the copper K electrons to calibrate the parts of the spectrograph in which the other, less deviable, groups were recorded This was done by making multiple exposures with a copper target, the current in the field coils being suitably altered at intervals during an exposure—the method being, in effect, an obvious modification of the "bracketing" technique employed by Aston with his mass-spectrograph.

We give below (table 1) the new results. rH is in absolute oersted cm. The equivalent frequency (ν^*/R) in Rydbergs is (purely to facilitate comparison with the earlier results) calculated from rH by using the following (in part superseded) values of the necessary constants:

$$e/m = 1{\cdot}757 \times 10^7\,\text{e.m.u. g.}^{-1}, \quad e/h = 7{\cdot}2827 \times 10^{16}\,\text{e s.u. erg}^{-1}\,\text{sec.}^{-1},$$

$$e/m \times e/h = 1{\cdot}2795 \times 10^{24}, \quad c = 2{\cdot}998 \times 10^{10}\,\text{cm. sec}^{-1}, \quad R = 109737{\cdot}4\,\text{cm.}^{-1}$$

The eleven values of (ν^*/R) in table 1 are to be compared with corresponding quantities obtained with molybdenum $K\alpha_1$ rays (Robinson and Clews 1935), and the first ten with corresponding values obtained with copper (Robinson 1934*b*) and chromium (Robinson 1934*a*) radiations. The primary X-ray frequencies (ν/R) on the Siegbahn scale are·

Ag $K\alpha_1$	1632·4
Mo $K\alpha_1$	1287·4
Cu $K\alpha_1$	592·7
Cr $K\alpha_1$	398·8

TABLE 1. PRIMARY X-RADIATION, AG $K\alpha_1$; FREQUENCY (ν/R) ON SIEGBAHN (CRYSTAL) SCALE, 1632·4 (ELG 1937)

Element	Level	rH	(ν^*/R)
79 Gold	M_{I}	466·1	1383·8
	M_{II}	469·5	1403·8
	M_{III}	474·7	1434·4
	M_{V}	481·0	1472·1
78 Platinum	M_{I}	467·5	1392·2
	M_{II}	471·0	1412·5
	M_{III}	$475{\cdot}8_5$	1441·3
	M_{V}	$482{\cdot}0_5$	1478·4
47 Silver	L_{I}	461·1	1354·8
	L_{III}	$466{\cdot}9_5$	1388·8
29 Copper	K	$389{\cdot}8_5$	973·3

The (ν^*/R) values are summarized in table 2

TABLE 2

Primary X-rays → level ↓	Ag $K\alpha_1$ (1)	Mo $K\alpha_1$ (2)	Cu $K\alpha_1$ (3)	Cr $K\alpha_1$ (4)
Au M_{I}	1383·8	1037·6	340·8	146·5
Au M_{II}	1403·8	1058·5	361·4	167·1
Au M_{III}	1434·4	1089·3	391·3	197·0
Au M_{V}	1472·1	1126·2	431·1	236·6
Pt M_{I}	1392·2	1046·4	350·7	156·3
Pt M_{II}	1412·5	1066·9	370·7	176·4
Pt M_{III}	1441·3	1095·7	398·8	204·4
Pt M_{V}	1478·4	1132·7	437·5	243·0
Ag L_{I}	1354·8	1009·6	312·7	118·6
Ag L_{III}	1388·8	1042·6	346·4	152·2
Cu K	973·3	627·3	—	-

Subtracting column 2 of table 2 from column 1, we get eleven values of (ν^*/R) (Ag–Mo) ranging from 345·1 to 346·2, with a mean of 345·7. The corresponding X-ray spectroscopic value is

$$(\nu/R)\ (\text{Ag–Mo}) = 1632{\cdot}4 - 1287{\cdot}4 = 345{\cdot}0.$$

From columns 3 and 1 we get ten differences ranging from 1040·9 to 1043·1, with a mean of 1042·1. The X-ray value is (1632·4 − 592·7) = 1039·7.

Columns 4 and 1 give ten values ranging from 1235·4 to 1237·4, mean 1236·4. The X-ray value is (1632·4 − 398·8) = 1233·6

The proportional differences in each case, 0·20, 0 23 and 0·23 % respectively, are identical to within the experimental error The mean percentage difference, 0·22, is a little higher than that previously adopted (0·17, Robinson 1936). It agrees very closely with a more recent value (0·23 %) of Robinson and Mayo (1939) The last-quoted value was not, however, obtained by the difference method, and the agreement must be regarded as accidental

The measurements reviewed above cover practically the whole range in which our magnetic spectrograph can be profitably used. It has already been pointed out that we cannot work with primary radiations appreciably harder than silver $K\alpha$, and at the other end of the scale there are obvious photographic and other difficulties in working with rays much softer than the chromium $K\alpha$. On the whole, there seems to be every reason to be satisfied with the internal consistency of our measurements—that of the X-ray data taken from Siegbahn and others may be accepted without question. It must not however be forgotten that as the whole of our measurements have been made in the same spectrograph, in as nearly as possible identical conditions, they may all be affected by a common systematic error. The only direct confirmation we have of our results is from the experiments of Kretschmar (1933) with molybdenum K rays. We have already shown (Robinson and Clews 1935) that, in spite of certain apparent discrepancies, Kretschmar's results are very probably in close numerical agreement with ours. There is, however, no independent confirmation of our work with chromium, copper and silver K radiations.

We take, as the mean of all our measurements, 0·21 % as the best value of the difference between the photoelectric (ν^*/R) and the conventional Siegbahn frequencies (ν/R), the (ν^*/R) being the larger by this amount This difference is increased by 0·20 % when the X-ray spectroscopic values are expressed in true Rydberg units ($\rho = 1{\cdot}0020$) In other words, according to our experiments the true value of the product $e/m \times e/h$ is less by 0·41 % than the value we adopted when calculating the (ν^*/R) values. That is, the value we find now for $e/m \times e/h$ is $0{\cdot}9959 \times 1{\cdot}2795 \times 10^{24} = 1{\cdot}2743 \times 10^{24}$ (or $3{\cdot}8201 \times 10^{34}$ if both charges are in e.s.u.). The value obtained by Robinson and Mayo (1939, p. 199) was $1\ 2741 \times 10^{24}$. Through a slip in entering the result, this was erroneously given as $1{\cdot}2749 \times 10^{24}$. We have to thank Professor R. T. Birge for drawing our attention to this slip, which, being purely an error in entering an antilogarithm, did not affect any other figure given in the paper.

Our new value of h/e, assuming $e/m = 1·759 \times 10^7$, is

$$\frac{1·759 \times 10^7}{1·2743 \times 10^{24}} = 1·380_3 \times 10^{-17},$$

and if we take e as $4·8022 \times 10^{-10}$, we have $h = 6·62_9 \times 10^{-27}$. This, as is well known, is appreciably greater than the value $(6·609 \times 10^{-27})$ given by recent accurate determinations of the short-wave limit of the continuous X-ray spectrum. On the other hand, it agrees well with the value deduced by Birge (1939) from the relation

$$h = \sqrt[3]{\frac{2\pi^2 e^5}{Rc^2 e/m}},$$

viz. $(6·6236_5 \pm 0·0024) \times 10^{-27}$.

The discrepancy between the two values of h has been fully discussed by Birge in a number of papers. In addition there are recent discussions by Darwin (1940), Du Mond (1939) and Dunnington (1939), to which we cannot usefully add anything. Our value of h/e, with $e = 4·8022 \times 10^{-10}$, gives $\frac{1}{137·1_4}$ for the fine-structure constant.

We have pleasure in acknowledging our indebtedness to the Government Grants Committee of the Royal Society for grants which provided for the greater part of the apparatus used in this work, and to other workers, in particular Professor R. T. Birge, for kindness in supplying information.

Note by H. R. R. In this, as in all the other photoelectric work from this Laboratory, the field current was measured potentiometrically, using a Weston cell and standard resistances which were tested at the National Physical Laboratory. The standardizations were given in terms of international electrical units, and a correction to absolute amperes was made when the magnetic fields were calculated. Unfortunately, I cannot now find my original notebook containing the details of the correcting factors used. From memory, I think that almost exactly 1 part in 10,000 was subtracted from the indicated currents in the field coils (or, more precisely, from the factor finally used as the effective constant of the coils). The published (ν^*/R) values have therefore already been reduced by about 1 part in 5000 on account of the difference between international and absolute amperes.

I regret that this was never mentioned in any of the papers; it certainly ought to have been stated explicitly somewhere, but in fact no real importance was attached to this correction. It was made purely as a matter of routine, and not from a belief that it was warranted by the accuracy of our experiments. A correction as small as this would in the circumstances almost certainly have been neglected if it had been at all difficult to apply.

REFERENCES

Birge, R. T. 1939 (August) Private communication.
Darwin, C. G. 1940 *Proc. Phys. Soc.* **52**, 202.
Du Mond, J. W. M. 1939 *Phys. Rev.* **56**, 153.
Dunnington, F. G. 1939 *Rev. Mod. Phys.* **11**. 65.
Elg, S. 1937 *Z. Phys.* **106**, 315
Kretschmar, G. G. 1933 *Phys. Rev.* **43**, 417.
Robinson, H. R. 1934*a* *Proc Phys. Soc.* **46**, 693.
— 1934*b* *Phil. Mag* (7), **18**, 1086.
— 1936 *Phil. Mag.* (7), **22**, 1129.
Robinson, H. R., Andrews, J. P. and Irons, E. J. 1933 *Proc. Roy. Soc.* A, **143**, 48.
Robinson, H. R. and Clews, C. J B. 1935 *Proc. Roy. Soc.* A, **149**, 587.
Robinson, H. R. and Mayo, R. L. 1939 *Proc Roy. Soc.* A, **173**, 192

The electric strength of some solid dielectrics*

BY A. E. W. AUSTEN AND S. WHITEHEAD

(*Communicated by N. F. Mott, F.R.S.—Received* 29 *April* 1940)

Methods are described by which the 'intrinsic' electric strength of solid dielectrics may be defined and evaluated. It is shown that the magnitudes of and the effect of temperature and thickness upon the electric strengths of certain crystals agree with Fröhlich's theory of electronic breakdown, as also does the effect of disordered structure and microstructure in similar instances. On the other hand, departures from theory occur with complex organic dielectrics and also with crystals when, with the latter, certain limits, e.g. of temperature, are exceeded.

1. INTRODUCTION

The paper describes some recent measurements of the intrinsic electric strength of certain solid dielectrics; a preliminary note of some of the results has already been published (cf. Austen and Hackett 1939). It was proposed to determine the actual values of electric strength of certain materials and the effect of temperature, thickness in the case of very thin layers, crystalline form and state of purity.

---

* Based on Report Reference L/T 114 of the British Electrical and Allied Industries Research Association.

The experiments were suggested by recent theoretical work on dielectric breakdown. Theories have been given by, among others, Zener (1924), von Hippel (1931), Fröhlich (1937) and Seeger and Teller (1939). Zener's theory, following earlier work by Whitehead and by Joffé, pictures a kind of pulling out of the electrons from the ions: it predicts a rapid increase of current in a dielectric as the field is raised, and in view of recent work must be considered to describe the electronic currents observed before breakdown, rather than the breakdown itself, which appears as a sudden instability. Von Hippel first suggested that intrinsic breakdown was of the nature of an electron avalanche, similar to that occurring in gases. Fröhlich and Seeger and Teller give mathematical developments of this idea. In a recent paper Fröhlich (1939*b*) has given reasons for believing that Seeger and Teller's results are incorrect, and since, moreover, the latter authors make no predictions as to the variation of electric strength with temperature or thickness, in the comparison of the present experiments with theory only Frohlich's work will be considered

Fröhlich's theory makes the following predictions

(1) The intrinsic strength should *decrease* as the temperature is lowered below the Debye characteristic temperature of the material, the effect being most marked for substances with low Debye temperatures.

(2) The strength should be independent of thickness for thick specimens, but should increase for very thin specimens of thickness of the order of a few mean free paths. (One mean free path is of order 10^{-6} cm., cf. Fröhlich and Mott 1939.)

(3) The strength of a mixed crystal is greater than the weighted mean of those of the pure components, and greater for a given substance in an amorphous than in a crystalline form. Impurities which form large aggregates, e.g. colloidal particles, may give a lower electric strength owing to distortion of the electric field.

(4) Microstructure due to differences in or relative displacement of neighbouring crystallites should probably not affect the electric strength.

Fröhlich's theory also gives a numerical formula from which the electric strength of any thick specimen of an ionic crystal can be calculated if the absorption spectrum of the crystal in the infra-red is known. The expression is

$$E = 5 \times 10^{4}\left(\frac{z}{V_r}\right)^{\frac{1}{2}} \lambda_0 \left\{\sum_{i=1}^{i=r} \frac{\epsilon_i - 1}{\lambda_i}\,\phi_i(T) \sum_{i=1}^{i=r} \frac{\epsilon_i - 1}{\lambda_i^2} \log Y_i\right\}^{\frac{1}{2}} \text{V/cm.}, \qquad (1)$$

where

V = molecular volume,

z = number of atoms per molecule,

$(\epsilon_i - 1) = 4\pi \times$ polarizability corresponding to the oscillation of residual wave-length λ_i,

$$\phi_i(T) = 1 + \frac{2}{e^{h\nu_i/kT} - 1},$$

T = abs. temp.,

$$Y_i = \frac{1}{\nu_i}\sqrt{\frac{J}{2m}}\left(\frac{2n}{r}\right)^{\frac{1}{2}},$$

J and λ_0 are the energy (in absolute units) and wave-length (in Angstroms) respectively corresponding to the first maximum of the ultra-violet absorption band,

n = number of atoms per cm ^{3},

r = number of residual ray frequencies,

m = mass of an electron.

If the approximation is made of a single mean residual ray frequency ν, equation (1) reduces to

$$E = 1{\cdot}64 \times 10^5 V^{-\frac{1}{2}} \lambda_0 \frac{(\epsilon - \epsilon_0)}{\lambda^{\frac{1}{2}}} \Big/ \sqrt{\left(1 + \frac{2}{e^{h\nu/kT} - 1}\right)} \text{ V/cm.}, \tag{2}$$

where

λ = residual ray wave-length, in Angstroms,

ϵ = static dielectric constant,

$\epsilon_0 = n^2$ = contribution of the electrons to the dielectric constant.

Von Hippel's experiments (1935) have shown that the theory gives values of electric strength of the alkali halides at room temperatures in satisfactory agreement with experiment, and also that the strength of mixed crystals is greater than that of either constituent. The present paper gives an account of work in which experimental methods suitable for a wide range of materials and conditions have been devised and describes experiments as follows:

The variation of the electric strength of potassium bromide with temperature has been measured.

The absolute value of the electric strength of mica, together with its variation with temperature has been determined.

Measurements of the electric strength of thin films of mica have been extended to the region where the strength increases considerably.

The strengths of fused and crystalline quartz are compared.

In addition the electric strengths of a coloured glass and the corresponding basic glass have been measured and organic dielectrics to which Fröhlich's theory does not apply investigated.

2. The nature and measurement of intrinsic breakdown

The electrical breakdown of a dielectric is frequently influenced by the thermal and special electrical conditions under which the test is made, as, for example in breakdown due to discharges in a surrounding liquid or gaseous medium or breakdown due to thermal instability. It is clear that for the present purpose the electric strength measured should be independent of such conditions and it has been found that a type of breakdown exists which is a property of the physical nature of the dielectric and its temperature only This has been called intrinsic breakdown and the corresponding field strength the intrinsic electric strength.

It is to be expected *a priori* that intrinsic breakdown should possess the following properties·

(*a*) The electric strength should be independent of thickness over a fairly wide range, except for very small thicknesses.

(*b*) The electric strength should be independent of the nature or duration of the electric stress provided no appreciable change of temperature occurs as a result of the application of the electric stress and provided the duration (e.g. with impulse voltage) is sufficient to permit the development of the discharge mechanism envisaged.

(*c*) The actual discharge should result in the local destruction of the structure and should occur wholly within the dielectric and in the region where the field is most intense.

The existence of an intrinsic electric strength postulates a homogeneous dielectric free from imperfections and weak spots. It is therefore preferable that as small an area of the dielectric as possible should be exposed to electric stress, but, at the same time the field must be calculable and sufficiently uniform in the direction of the field to avoid incomplete breakdown. For these reasons two spherical electrodes (or one sphere and a plane) were used, where, provided the radius of the sphere is large compared with the thickness of the dielectric, the intense portion of the field at the sphere vertex is substantially uniform through the thickness of the dielectric but is limited to a small area in the neighbourhood of the vertex.

Materials are frequently available in the form of thin plane sheets and if they are tested between spheres the medium between the spherical electrode and the dielectric, in the vicinity of the point of contact, is stressed to a value comparable with the stress in the test specimen. No gaseous or liquid dielectrics in which the test may be carried out have strengths as high as that of some solid dielectrics and the immersion medium then frequently breaks down first and the resulting discharge, either by chemical action or by causing local concentrations of stress, causes premature breakdown of the test specimen. Numerous attempts to overcome this difficulty have been made by immersing the system in liquids of high electric strength or in waxes and sometimes subjecting the whole to high pressures still further to increase the electric strength.

In tests on solid dielectrics of high strength it is necessary to select an immersion medium not only of high strength but also of suitable resistivity, in the case of D.C tests, or specific impedance for A.C. tests, with respect to that of the specimen. If σ, ϵ, δ and E are the conductivity, permittivity, loss angle and electric strength respectively, the subscript S referring to the dielectric and M to the immersion medium, then

$$\sigma_M E_M > \sigma_S E_S \text{ for D.C.}$$

or

$$E_M \epsilon_M \sec \delta_M > E_S \epsilon_S \sec \delta_S \text{ for A.C.}$$

are rough conditions for the avoidance of discharges in the medium. These conditions are not precise since they neglect tangential stress and they are difficult to apply on account of the variation of the quantities with field strength. Theoretically the occurrence of breakdown at the point of contact of the spherical electrode is evidence of a satisfactory test, but the point of contact is difficult to locate with precision and there may be a thin film of the immersion medium so that a breakdown apparently at the point of contact may be spurious On the other hand, with semi-conducting media the specimen is stressed to a value closely approximating to the maximum value at the point of contact over an appreciable area so that a breakdown appreciably separated from the point of contact is not necessarily spurious. A better criterion is the observation of the apparent electric strength for variation of the immersion medium, and if the apparent electric strength is not decreased by a considerable decrease of $\sigma_M E_M$ or $E_M \epsilon_M \sec \delta_M$ it is fairly certain that the breakdown stress is reached in the specimen before the immersion medium is overstressed.

While it is not difficult to find liquid immersion media suitable for use at room temperatures, e.g. mixing castor oil with nitrobenzene permits a wide variation of conductivity without much change of electric strength, at high or low temperatures few of the normal dielectric liquids are available and any such control of conductivity is impossible. It is then desirable to adopt a specimen shaped to the contour of the electrodes as shown in figure 1 A, the electrodes, e.g. in the form of colloidal graphite, being deposited on the dielectric itself. The system then consists of test dielectric and electrodes only, so that provided breakdown takes place within the electrodes, there can be no effect due to external discharges. The lateral dimensions of the specimen and its thickness outside the deposited electrodes necessary to prevent flash-over or breakdown outside the electrodes can be reduced by immersing in a suitable dielectric liquid. Such a liquid will be called an ambient medium as distinct from the immersion medium of the plane sheet tests.

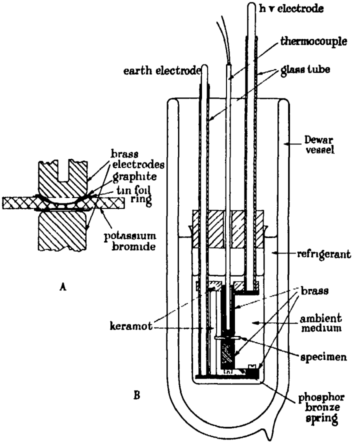


FIGURE 1. Specimen (A) and apparatus (B) for measuring the electric strength of potassium bromide.

3. Results

(a) *Potassium bromide*

The material was obtained in the form of an artificially prepared single crystal of cylindrical shape about 4 cm in diameter and 4 cm. long. It was found possible to split complete circular plates down to a thickness of 0·5 cm., while plates of linear dimensions about 1·5 cm. could be obtained down to a thickness of about 0·15 cm. It appeared, therefore, to be necessary to resort to grinding to obtain specimens of thickness suitable for electric strength measurements, i.e. about 0·02 cm. Since it was desired to make measurements at low temperatures it was decided to attempt to prepare specimens with a spherical recess.

Tinfoil rings were cemented to the specimen at the edge of the recess and the specimen painted with a suspension of colloidal graphite in acetone to form the electrode proper Contact with the tinfoil rings was made by two cylindrical brass electrodes with rounded ends and maintained by spring pressure, as shown in figure 1 B. It was found that direct immersion in a refrigerant such as liquid nitrogen caused cracking of the specimen and the electrode system was therefore placed in a liquid, petroleum ether at low temperatures or transformer oil at room or higher temperatures, and the whole assembly placed in a refrigerant in a Dewar vessel or a heated oil bath for tests at low or high temperatures respectively. The temperature was measured by means of a thermocouple in a thin-walled glass tube sunk into the high-voltage electrode and withdrawn just before the application of voltage

Tests were made at room temperature on specimens (*a*) with a recess dissolved out by means of a jet of water and (*b*) with the recess and the plane side ground with carborundum powder (no. 600) in oil, on a copper tool of radius of curvature 1 cm , and polished with rouge or polishing alumina on a chamois leather-covered tool; the values obtained were both inconsistent and rather low, i.e. about $0{\cdot}5 \times 10^6$ V/cm., compared with the value obtained by von Hippel of $0{\cdot}7 \times 10^6$ V/cm. It was found, however, that remarkably consistent values of about $0{\cdot}79 \times 10^6$ V/cm. were obtained with specimens ground with no. 600 carborundum but not polished, the surface of the material then having a fine matt appearance, the irregularities presumably being of linear dimensions of the same order as those of the carborundum grains, i.e. about 0·001 cm. Microscopic examination of the edge of a cracked specimen, however, showed that the depth of the irregularities was considerably less. The electric strength showed no variation with thickness over a range from 0·015 to 0·04 cm.

at room temperature, and measurements were made at lower and higher temperatures with ground but unpolished specimens of thickness of order 0·02 cm. The results are compared with the calculated values of Fröhlich in figure 2.

The following check tests were made. Two specimens were treated as for a test at low temperature but allowed to warm up to room temperature and then tested: normal values were obtained indicating that the specimens did not suffer permanent damage on cooling. The possibility of water absorption was also considered and specimens were prepared, graphited,

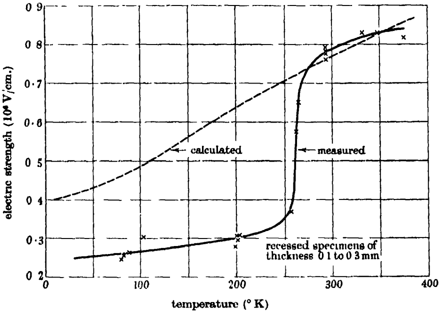


FIGURE 2. Electric strength of potassium bromide.

baked at 115° C for 2 hr., cooled over phosphorus pentoxide and tested in petroleum ether containing phosphorus pentoxide, both at room and low temperatures. Normal values were obtained.

It will be noted that the direction of the variation of electric strength with temperature is in agreement with theory but that the decrease is markedly greater than that obtained theoretically. Von Hippel (1939) has since published the results of tests on potassium bromide and obtains a similar fairly sudden decrease of strength with decrease of temperature, which, while also greater than that required by the theory, is not as great as that observed by the present authors. A complete account of von

Hippel's experimental technique, however, has not yet been given, and in particular the thickness of the specimens is not stated so that a discussion of the discrepancy is not possible.

The remaining possible source of error in the authors' experiments would appear to be the nature of the surface of the specimens. It was observed, however, that the surface of the polished specimens was not free from defects, nor was it found possible to obtain a perfect surface. Tests on ground specimens of crystalline quartz, however, gave the same value as polished specimens so that it would appear that the low and erratic values for polished potassium bromide are due to surface defects and that the values for ground specimens are valid.

(*b*) *Mica*

The considerable body of research conducted by the Electrical Research Association on the electric strength of mica for industrial purposes of thicknesses available commercially, i e. greater than 0·002 cm., is being published elsewhere and only those parts of theoretical significance will be described here. Tests were made on plane sheets with spherical metal electrodes and immersion media, as already described, and on specimens with a spherical recess in a manner essentially similar to that used for potassium bromide. Further description of experimental methods is therefore unnecessary and only the results will be given

Measurements at room temperature on plane specimens using as immersion medium a mixture of castor oil and nitrobenzene (resistivity about 2×10^8 ohm-cm. electric strength about $0{\cdot}1 \times 10^6$ V/cm.) showed that there was no variation of strength with thickness in the range 0·002–0·007 cm., and that the D C. and 50 cyc. A.C. crest values were the same. Impulse measurements were made at the National Physical Laboratory in the course of another investigation (Davis and Lacey 1939). Values obtained using a $1/5\,\mu$sec. standard wave for plane sheet specimens are compared with D.C. values for specimens from the same consignment in table 1, which shows that although the impulse values are usually higher, the difference is not large compared with the experimental dispersion, which tended to be high in the impulse tests, where distilled water was used as an immersion medium.

X-ray measurements also on the same consignment were made at the National Physical Laboratory. Table 1 shows the lattice spacing and the length of the diffraction arc which is a measure of the mean tilting of the axes of individual crystallites about the mean direction. It will be seen that although the thinner specimens showed a more pronounced microstructure, i.e. a greater relative tilting of the individual crystallites, their electric strengths were, if anything, higher. This supports Fröhlich's view that the intrinsic electric strength is not "structure sensitive" in the commonly accepted meaning of the term.

TABLE 1 EFFECT OF DURATION OF STRESS AND MICROSTRUCTURE FOR MUSCOVITE (RUBY) MICA

Type	Nominal thickness mm.	Electric strength (D.C.) in 10^6 V/cm.		Ratio impulse to D.C.		Length of diffraction arc* cm.	Lattice spacings perpendicular to cleavage plane (A)
		Mean	Maximum	Mean	Maximum		
Clear	0·02	10·15	10·7	1·28	1·26	3·25	9·946
	0·04	10·3	10·8	1·16	1·25	3·00	9·946
	0·06	9·75	10·0	1·08	1·50	2·60	9·946
Stained	0·02	11·6	12·0	0·99	1·17	3·20	9·957
	0·04	10·5	11·5	1·05	1·17	3·00	9·958
	0·06	10·6	11·5	0·80	1·00	2·65	9·958

* Length of recorded X-ray reflexion measured along the diffraction ring. This is a measure of the relative tilting of the planes of neighbouring crystallites.

Owing to lack of the requisite optical data, it is not possible to calculate accurately the electric strength of mica. Fröhlich, however (cf. Fröhlich 1939*a*), assuming that the effective residual wave-lengths are the same as those of quartz, i.e. $\lambda_1 = 8{\cdot}4\mu$, $\lambda_2 = 20{\cdot}5\mu$ and that $\epsilon_1 = \epsilon_2$, and that the wave-lengths of the first ultra-violet maxima λ_0 are the same, obtains a value between $6{\cdot}6 \times 10^6$ and $7{\cdot}5 \times 10^6$ V/cm., while the assumption that $\lambda_1 = 8{\cdot}4\mu$ is the only important residual wave-length gives a strength of 15×10^6 V/cm. which is an upper limit to the theoretically possible value. The variation with temperature also depends on the residual wave-lengths, and the curves plotted in figure 3 are obtained by assuming the observed value at room temperature and residual wave-lengths of 10 and 20μ respectively. It will be seen that the theoretical curve for 10μ and the maximum experimental values are in satisfactory agreement up to 450° K. At higher temperatures the electric strength decreases greatly, as is common with solid dielectrics, since the structure itself frequently changes, other breakdown mechanisms intervene and Fröhlich's theory no longer applies.

(c) *Thin layers of mica*

Joffé and his collaborators (cf. Semenoff and Walther 1928) claimed to have observed electric strengths up to 100×10^6 V/cm. for glass and mica of thickness of order 10^{-5} cm. These values, however, were subsequently withdrawn (Joffé and Alexandrow 1932), and in a new investigation in which errors due to a potential drop in the graphite electrodes and in the measurement of thickness were eliminated, it was found that the electric strength of mica increased only slightly from $9{\cdot}4 \times 10^6$ V/cm. at a thickness of 5×10^{-3} cm. to 12×10^6 V/cm. at $0{\cdot}7 \times 10^{-4}$ cm. The present measurements extend this range to still smaller thicknesses.

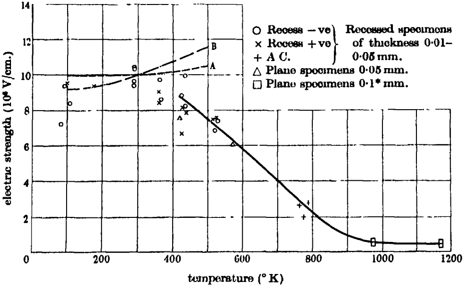


FIGURE 3. Electric strength of clear ruby muscovite mica. A, Theoretical curve for mean residual ray of wave-length 10μ; B, Similar curve for wave-length 20μ.

* At the lower temperatures breakdown of specimens of this thickness occurred at relatively low stresses as a result of breakdown of the immersion medium.

It was found possible to obtain sheets of mica down to thicknesses of order 0·001 cm. by splitting from thicker material. If these sheets were torn, small areas at the edges of the tear, of linear dimensions of a few millimetres, were sufficiently thin to show the colours of thin films, i.e. thickness of order 5×10^{-5} cm These pieces were frequently preserved, if, after beginning detachment by tearing, the process was completed by sliding in the edge of a piece of paper, and in this way specimens of a range of thickness down to 2×10^{-5} cm. were obtained, together with a few pieces too thin to show the colours of thin films, i.e. less than 10^{-5} cm.

The thickness of the specimens were determined approximately by visual observation in diffused daylight of the colours corresponding to normal and grazing emergence for both reflected and transmitted light.

The corresponding wave-lengths were estimated and it was then generally possible to deduce the orders of the various interference colours and therefore the thickness. The order was confirmed, or if necessary, one of two possible values selected by examining in the approximately monochromatic light from a sodium flame, viewing the light transmitted normally and observing the changes of intensity on rotating the specimen to the position for grazing emergence.

The thicknesses were measured accurately (after the breakdown measurements, except in the case of two specimens less than 10^{-5} cm. thick) at the National Physical Laboratory, by an optical method using a Fizeau interferometer. The average accuracy obtained was $\pm 3 \times 10^{-8}$ cm., i.e. $\pm 3\%$ for the thinnest specimens, and one fairly large specimen was also measured by a mechanical method, the value of $6{\cdot}10 \times 10^{-5}$ cm. being in good agreement with the value of $6{\cdot}09 \times 10^{-5}$ given by the optical method. The thickness estimated from the colour of the specimens was in fair agreement with the optical measurements, the error being generally less than 10%.

Breakdown tests were made between an approximately plane electrode of polished steel mounted on the stage of a metallurgical microscope and a polished steel sphere of diameter 0·3 cm. on the lower end of a short, fairly stiff helical spring inserted in place of one objective in a double nosepiece. The spherical electrode could be raised by the microscope focusing movement and replaced by a 1 in. objective, which, together with a micrometer eyepiece giving a magnification of 40 diameters, made possible the location of the point of contact to an accuracy of 0·005 cm. which was adequate for selecting a suitable part of the specimen.

A D.C. voltage was applied at a rate such that breakdown occurred in about 30 sec. and the voltage was measured directly by means of an electrostatic voltmeter. It was expected that at the comparatively low voltages necessary to cause breakdown, discharges in the immersion medium, if they occurred at all, would be without effect on the breakdown voltage. This view was confirmed by breakdown measurements on one of the thicker specimens in air and also immersed in a mixture of castor oil and nitrobenzene of resistivity approximately 2×10^{8} ohm-cm., which had been shown to be a suitable medium in tests on thicker specimens. The value obtained was the same in both cases and the tests were therefore made in air.

A number of tests were made on each specimen and in most cases at least five values were obtained: the results are plotted in figure 4. The greatest difference from the mean value for any one specimen was 15%, while the dispersion was generally of order ± 8 % and satisfactory measurements were made down to a thickness of about 2×10^{-5} cm. Two thinner specimens of thickness about 10^{-5} cm. were measured at the National Physical Laboratory, but considerable difficulty was encountered in manipulating them owing to their extreme fragility and tendency to adhere to any clean metal surface. One was broken into fragments too small for breakdown tests in placing it on the plane electrode and although a number

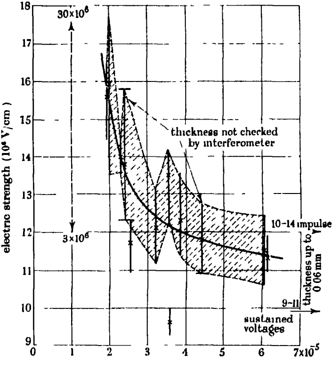


FIGURE 4. Electric strength of clear ruby muscovite mica. × Indicates mean value. Vertical lines indicate range of dispersion.

of tests were made on the other, the breakdown voltage showed a very wide dispersion from 30 to 290 V and the measurements cannot be regarded as satisfactory.

Fröhlich (as yet unpublished) has examined quantitatively the effect of thickness, concluding for mica that the mean free path is of order 5×10^{-7} cm. and that the electric strength should increase by 60 % at a thickness of order 5×10^{-5} cm. and by 100 % at 2×10^{-5} cm. Figure 4 shows that the experiments are in satisfactory agreement with this result.

(*d*) *Quartz*

Quartz and fused silica (0·05–0·1 mm. thick) were tested at room temperatures in a manner similar to potassium bromide. The observed value for quartz with the direction of the electric stress parallel to the optic axis was $6{\cdot}7 \times 10^6$ V/cm. and the value calculated by Fröhlich $3{\cdot}3 \times 10^6$ V/cm. with an upper limit of $7{\cdot}4 \times 10^6$ V/cm., a reasonably good agreement. According to the theory a disordered structure should increase the electric strength E, but allowance must be made for the decrease in the dielectric constant ϵ when quartz is fused, from 4·6* to 3·9.* This may be corrected by the approximate formula†

$$E \propto \epsilon - 2{\cdot}3.$$

The electric strength of a hypothetical quartz having the same ϵ as the fused silica should therefore be $4{\cdot}7 \times 10^6$ V/cm. (from the experimental results on quartz). The measured electric strength of fused silica is $5{\cdot}4 \times 10^6$ V/cm. so that the disordered structure seems to increase the electric strength.

(*e*) *Glass*

If a conducting particle is placed in a dielectric subjected to an otherwise uniform electric field, the field strength in the vicinity of the particle is increased, and it was therefore expected that the electric strength of a dielectric containing metallic particles would be lower than that of the same dielectric without particles. It was proposed to test this hypothesis by making tests on a glass the colour of which was due to the presence of colloidal metallic particles and on the corresponding basic glass, and also to verify that the increase of electric strength observed by von Hippel‡ for mixed crystals applied also to a glass containing a metal in true solid solution.

Samples of gold ruby glass, the colour of which is supposed to be due to the presence of colloidal particles of metallic gold, and a basic lead glass of identical composition§ but without gold were specially prepared in the Department of Glass Technology, Sheffield University, by Professor Turner and Dr Seddon.

Tests made on ground specimens of the type used for potassium bromide gave unsatisfactory results which were attributed to flaws in the glass.

---

* These values were determined for the actual materials from which the specimens were made.

† See Fröhlich (1939).

‡ Von Hippel (1935).

§ Composition (calculated from ingredients) was SiO_2 47·36 %, PbO 39·78 %, K_2O 10·05 %, Sb_2O_3 1·93 %, Fe_2O_3 0·10 %, MnO 0·79 %. The gold was estimated at 0·04 % in the coloured glass.

Thin sheets were obtained by blowing on the end of a quartz tube in an electric furnace and satisfactory tests were made on specimens of thickness from 3×10^{-4} cm. to 15×10^{-4} cm. between sphere and plane electrodes in an immersion medium of resistivity 10^8 ohm-cm. The thickness was measured *in situ* on the dial micrometer which was used for supporting the electrodes. The dispersion was rather large, the mean of all values for the basic glass being $5\ 2 \times 10^{-6}$ V/cm. and the mean of the maximum values for each specimen $6{\cdot}0 \times 10^6$. The corresponding results for the basic glass were $5{\cdot}0 \times 10^6$ and $6{\cdot}0 \times 10^6$ V/cm. respectively.

This value is much higher than those previously obtained, e.g. it is about double the value given by Moon and Norcross (1929), and the discrepancy is probably due to the unsuitable electrode systems and/or immersion media used hitherto.

The failure of the gold to produce any measurable decrease of electric strength is probably due to the small size of the colloidal particles. It is improbable that they are of linear dimensions greater than 2×10^{-5} cm., and if they are assumed to be spherical, although the mean field E would be increased to $3E$ at the surface of each sphere, the increased field is maintained only over a distance comparable with the radius of the particle. thus at a distance of 10^{-5} cm., the field would not exceed $1{\cdot}25\ E$. The theory and experimental results on thin layers of mica indicate that breakdown can occur in an electron path of this length only at considerably increased fields.

In view of this negative result and the difficulty of making reliable measurements on glasses made in small batches it was decided not to examine coloured glasses of the solid solution type. The change of electric strength would, in any event, be small since the amount of metal which may be introduced is minute.

(*f*) *Organic dielectrics*

A brief note is given of the behaviour of more complex organic dielectrics as employed for industrial electrical insulation which show properties different from those of the preceding materials. The tests to be quoted were made on synthetic resin (Bakelite), shellac, bituminous and natural resin oil varnishes, the last two containing, however, a small addition of synthetic resin. Films of these varnishes were cast on tin foil on which small spherical projections had previously been formed. After appropriate heat treatment the tin foil was dissolved in mercury leaving a film of varnish containing a number of spherical recesses and forming thus a multiple specimen of the type already described. The recesses were sprayed on both sides with colloidal graphite to form the conducting coatings. The samples were not immersed directly in liquid nitrogen since the rapid cooling caused the films to crack, but were tested by a method similar to that employed for mica, using iso-pentane as ambient medium at low temperatures, and transformer oil at other temperatures.

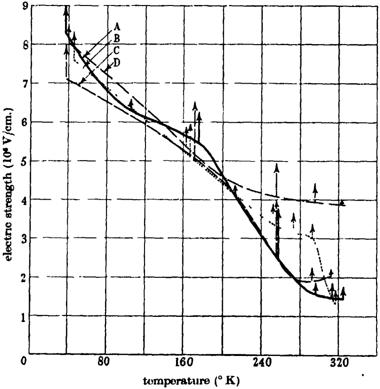


FIGURE 5. Electric strength of varnish films. A, clear baking oil varnish (0·03–0·05 mm.). B, shellac (dewaxed lemon) varnish (0·005–0·05 mm.) C, bituminous baking varnish (0·03–0·05 mm.). D, bakelite varnish (0·03–0·05 mm.). Curves are drawn through mean values. Vertical lines represent dispersion from mean to maximum.

The results are shown in figure 5 and indicate that the electric strength of these dielectrics decreases very greatly but in a somewhat irregular manner with increase of temperature. It is noticeable that at 80° K the electric strength is as high as that of some micas although much less at room temperatures. It is difficult to explain such an effect in any simple way on the basis of the theory of electronic breakdown as advanced and speculation must be deferred until pure organic dielectrics of known simple structure have been studied.

Acknowledgements are due to the Director of the Electrical Research Association for permission to publish this paper, to Dr Bragg (then Director) and members of the staff of the National Physical Laboratory for checking the thickness of the thin mica specimens and for the impulse voltage and X-ray data on mica extracted from National Physical Laboratory reports; to Professor Mott and Dr Fröhlich for their advice and for their good offices in obtaining the potassium bromide crystal; to Messrs Thermal Syndicate, Ltd., for the specimens of fused quartz.

The authors also wish to express their appreciation of the assistance of their colleagues Mr A. M. Thomas, Dr W. Hackett and Miss M. V. Griffith.

SUMMARY

Intrinsic electric strength, a property of homogeneous solid dielectrics, is defined and methods of measurement suitable for a wide range of materials and conditions are described.

A comparison of experimental results with Fröhlich's theory yields the following conclusions.

(*a*) The electric strengths of mica, quartz, potassium bromide and (from von Hippel's work) other alkali halides agree with the theory, though the limits of uncertainty due to lack of infra-red data for the first two are rather wide.

(*b*) The effect of temperature on the electric strengths of mica and potassium bromide agrees qualitatively with the theory, while for mica the agreement is also satisfactory quantitatively.

(*c*) The electric strength of mica increases when the thickness is reduced to the order of a few mean free paths and the agreement is quantitative to the accuracy with which calculation is possible. Indirect corroboration of the existence of this effect is afforded also by the absence of effect of colloidal metallic particles in glass.

(*d*) Fused quartz has a relatively higher electric strength than crystalline quartz corresponding to the prediction of the theory that disorder increases the electric strength. Experiments on mica indicate that, as the theory suggests, the electric strength is not sensitive to microstructure.

In addition the electric strength of mica decreases above 400° K, and complex organic dielectrics show a decrease of strength with increase of temperature; Fröhlich's theory does not apply in these cases.

REFERENCES

Austen and Hackett 1939 *Nature, Lond.*, **143**, 637.
Davis and Lacey 1939 *Conférence Internationale des Grands Réseaux Électrique à Haute Tension*, no. 304 Paris.
Frohlich 1937 *Proc. Roy. Soc.* A, **160**, 230.
— 1939*a* *Proc. Roy. Soc.* A, **172**, 94.
— 1939*b* *Phys. Rev.* **56**, 349.
Fröhlich and Mott 1939 *Proc. Roy. Soc.* A, **171**, 496.
v. Hippel 1931 *Z. Phys.* **67**, 707.
— 1933 *Z. Phys.* **68**, 309.
— 1935 *Ergebn. exakt. Naturw.* **14**, 79.
— 1939 *Phys. Rev.* **56**, 941.
Joffé and Alexandrow 1932 *Phys Z. Sowjet.* **2**, 527
Moon and Norcross 1929 *J. Franklin Inst* **208**, 705.
Seeger and Teller 1939 *Phys. Rev.* **54**, 515.
Semenoff and Walther 1928 *Die Physikalischen Grundlagen der electrischen Festigkeitslehre*. Berlin.
Zener 1924 *Proc Roy. Soc.* A, **145**, 523.

Melting and crystal structure

BY J. W. H. OLDHAM AND A. R UBBELOHDE

The Davy Faraday Laboratory, The Royal Institution

(*Communicated by Sir William Bragg, P.R S.—Received* 27 *February* 1940)

The role of various types of lattice flaws is discussed in their bearing on melting and crystal structure. Freezing-point determinations on lattices of a single molecule indicate the presence of such flaws in the crystals of polymethylene compounds. Freezing-point determinations on composite lattices of ketone + paraffin and ketone + ketone emphasize the importance of the rotation of the CO group for melting, and indicate that lattice holes are easily formed, without marked lowering of the freezing-point. Ordered composite lattices are formed when the number of holes does not exceed 12 % of all the methylene groups.

The results emphasize the importance of co-operative effects in lattice flaws. In the equilibrium state of a crystal, such flaws might arise both from energy effects at 0° K, expecially in metals, and from energy and entropy effects at higher temperatures. A picture of melting is suggested in which a network of co-operative flaws breaks up the crystal into a mosaic, or aggregate of crystalline micelles, and in which the solid-liquid transition is comparable with a gel-sol transition

Two main approaches have been explored in the discussion of melting. As the temperature of a crystal is raised, the heat motions of the molecules increase. In principle, it must be possible to explain the break up of a crystal to form liquid, in terms of these heat motions. Such an explanation emphasizes the mechanical aspect of melting, and defines the melting point as the temperature above which the solid no longer exhibits elastic resistance to a shearing stress (cf. Born 1939; Lucas 1938).

Mechanical theories of melting are of particular interest in linking up the phenomenon with that of plastic flow in crystals. With the experimental evidence at present available, however, the rather different approach from the standpoint of formal thermodynamics promises to give more detailed information on the relation between melting and crystal structure. On changing from solid to liquid, an assemblage of molecules increases its heat content, owing to work done against intermolecular attractions, and also increases its entropy, owing to increased freedom of motion in the liquid. At the melting-point, the free energy $G = H - TS$ of the liquid becomes equal to that of the solid. X-rays and other methods of investigation (e g. Debye 1939) indicate that both solid and liquid possess 'structure', and theories of melting aim at calculations of the heat content and entropy in terms of this structure (e.g. Lennard Jones and Devonshire 1939, Frank 1939; Irany 1939; Wannier 1939; Mott and Gurney 1939; Frenkel 1939).

The present paper describes further experiments designed to throw light on the relation between structure and melting in polymethylene lattices. As far as is known, the melting-points of long-chain paraffins and of their derivatives reach limiting values (convergence temperatures) as the chain length increases. The entropy and heat of melting appear to increase indefinitely with chain length (cf. Ubbelohde 1939). Since chemical means are available for introducing small perturbations in the lattice, effects can be investigated which might be much more difficult to detect in other compounds. For example, the data for the specific heat of octadecane (Ubbelohde 1938, referred to as (I)) suggested that a significant proportion of flaws might persist in the crystal in equilibrium, even below the melting-point. This would imply that the mechanism of melting involves the dissemination of crystal flaws to such an extent that the lattice can no longer maintain any kind of rigidity. Furthermore, the effects which result from the introduction of sheets of dipoles in the lattice (Oldham and Ubbelohde 1939*b*, referred to as (II)) suggested that torsional oscillations of the methylene chains were a factor of some importance in determining melting. The investigations discussed below give fresh experimental support for the process of melting suggested in the earlier papers. They include detailed measurements of cryoscopic heats of fusion over a narrow range of temperatures, and establish the existence of (ordered) solid solutions of a new type, whose structure throws fresh light on the melting of polymethylene lattices.

CRYOSCOPIC HEATS OF FUSION OF OCTADECANE AND DIOCTYL KETONE

If the number of flaws in a solid increases rapidly towards the melting-point, the lattice energy of the crystal must show a corresponding decrease. It might be expected that the liquid phase would not show changes of the same order of magnitude, and in this case there should be an appreciable drift in the heat change from solid to liquid, as the temperature of separation of the solid is lowered by the addition of impurities. The specific heat data for octadecane (I) suggest an increase in the heat of fusion of about 8 % over a few degrees fall in freezing-point. This should be detectable in careful determinations of the freezing-point depressions.

The effect of foreign molecules on this method of detecting premelting may be contrasted with their effect on specific-heat determinations. Foreign molecules may be distinguished according to whether they dissolve only in the liquid phase (Raoult impurities), or whether they dissolve fairly freely in the crystal as well as in the liquid phase. In specific-heat determinations, molecules dissolving freely in the crystal have only an indirect effect on premelting, whereas Raoult impurities can give rise to spurious specific heats owing to the production of liquid below the apparent melting-point. [Owing to a numerical error, the possible effect of Raoult impurities was underestimated in (I); the figure of 11 moles % on p. 298 should have been 1·3 mole %. The other reasons for believing that the specific-heat curves cannot be accounted for on the sole assumption of Raoult impurities are not affected by this mistake. A further point not previously mentioned in this connexion is that no two premelting curves can intersect if they are due to Raoult impurities, as can be seen by substituting different values of ΔT_M in the equation $\Delta C = L_f \Delta T_M / \Delta T^2$.]

Determinations of the cryoscopic constant are a useful check on the specific-heat curves, since the importance of different types of foreign molecules is reversed. Small amounts of Raoult impurities initially present will not affect the slope of the freezing-point curve, and are immaterial. Impurities dissolving freely in the crystal are more serious, since it was found (II, p. 332) that they lessen the apparent depression of the freezing-point, and might mask a drift in the cryoscopic constant. From the experiments described below, it seems likely that this type of impurity packs into the solid leaving holes, which dissolve some of the solute and thus lessen the depression of the freezing-point (equation (9)). In order to obtain significant values of ΔH_f attempts were made to keep the concentration $[N_2]$ of solute molecules in the crystals as low as possible, by carefully purifying the octadecane and dioctyl ketone, and by selecting as solutes molecules of such a shape as to make dissolution in the solid unlikely. In order to avoid uncertainties due to association, only hydrocarbons were used as solutes.

The equilibrium temperatures were measured on cooling the melts, using a calibrated micro-Beckmann thermometer. Reproducible results were obtained when the vessel containing the melt was insulated by an air gap from an outer bath kept within about 0·3° of the temperature of the melt, though no change in freezing-point could be detected when this interval was increased at least to 1° C. The melt was stirred by a platinum ring actuated by a hollow glass rod, and it was found important to shield it from sunlight, or even from powerful electric lights, which appeared to have a marked effect in raising the temperature of polymethylene liquids. Addition of the solute was made in the form of small weighed pellets.

The values of the cryoscopic constant are calculated using the Hildebrand equation

$$\log N_1 = \frac{\Delta H_f}{4{\cdot}575}\left(\frac{1}{T_M} - \frac{1}{T}\right), \tag{1}$$

and applying the method of least squares. The weights of solute added were chosen so as to give successive depressions of about 0·030° in the freezing-point of the melt, so that each temperature range corresponds with a Hildebrand equation smoothed from ten to twelve successive observations of the freezing-point (table 1).

TABLE 1. CRYOSCOPIC CONSTANTS

Dioctyl ketone $(C_8H_{17})_2CO$. Freezing-point 50·230 ± 0·005°. Solute: dibenzyl

Temp. range ° C	(kcal./mole)	(cal./mol./deg.)
50·230–49·925	13·30	41·15
49·907–49·597	13·19	40·85
49·285–48·916	13·77	42·74
48·876–48·436	13·78	42·83

The values may be compared with the average heat of fusion determined by a rougher method, over the range 50·0–47·0°, i.e. 13·7 kcal./mole (II, p. 331). Although a drift in the cryoscopic constant is suggested by these values, the effect is hardly sufficient to warrant an extension of the very laborious observations to other solutes. Premelting in octadecane might be expected to be considerably larger, owing to the fact that the energy required to produce a flaw of a given type is considerably smaller in a hydrocarbon (see table 2).

TABLE 2. CRYOSCOPIC CONSTANTS

Octadecane $C_{18}H_{38}$ Freezing-point 28 096 ± 0 005°.
Solute: napthalene (melting-point 81°)

Temp. range ° C	(kcal /mole)	(cal /mol /deg.)
28·096–27·817	13 98	46·59
27 764–27·528	14·35	47·71
27·493–27·084	15·12	50 33
27 020–26·718	15 21	50·71
26 659–26·360	15·13	50 51
26·208 25 949	15 08	50 40

A confirmatory series of experiments was carried out using camphene as solute. These gave evidence of a corresponding drift, but as the solute was of doubtful purity the numerical values are not given.

In the experiments with naphthalene, a second crystal form of octadecane was twice observed to separate from the melt Since the freezing-point was lower, this corresponds with a metastable phase which is probably identical with that previously described (Smith 1932; (I), plate 294). A rough cryoscopic estimate of the heat of fusion gave 13·5 kcal./mol.

The interpretation of the above results depends on the validity of the Hildebrand equation, and of Raoult's law on which it is based. Although the range of application of Raoult's law has been criticized (Guggenheim 1937), the good agreement with equation (1) obtained for certain pairs of molecules over a wide range of concentrations (see below, and cf. Hildebrand and Sweeny 1939) suggests that the observed drift in the heat of fusion is not due to a wrong application of Hildebrand's equation. The experiments are most simply interpreted on the view that a polymethylene lattice has an appreciable proportion of flaws in the solid state, in the neighbourhood of the melting-point. They do not, however, give any precise picture of what these flaws are like.

DEFECT LATTICES AND MELTING

Simple types of lattice flaws have been suggested in the case of regular lattices of spherically symmetrical molecules (Wagner and Schottky 1930). The heat content of the crystals increases when some of the molecules are removed from their equilibrium positions, and are either built on to the outside of the crystal, leaving 'holes' in the lattice, or are moved to interstitial positions, forming regions in the lattice where the repulsion energy is larger than the average. Diagrams illustrating these flaws are given in figure 11*a*, *b*. Owing to the fact that the crystal increases its entropy by the lack of order introduced, both types of flaw become more frequent as the temperature rises.

When there is no interaction between neighbouring flaws, a simple expression for the fraction n/N of defect sites in equilibrium at any one temperature is $n/N = e^{-E/2kT}$ (Jost 1933), where E is the 'hole energy'. In a polymethylene lattice the energy per CH_2 group is about $1\cdot 1 \times 10^{-13}$ erg, and if it were possible to produce holes with this energy in a crystal of a normal octadecane as much as 30 % of the lattice might be defective near the melting-point. It seems likely, however, that owing to the interaction of neighbouring sites in a crystal, 'co-operative' defects may be at least of equal importance for melting, as isolated flaws. A more precise picture of such flaws is discussed below (cf. figure 11*c*).

In order to test some of the possibilities experimentally, it was decided to investigate the melting of solid solutions of closely related molecules.

The melting of solid solutions An exhaustive discussion of the formal thermodynamics of solid solutions, in equilibrium with their melts, would occupy undue space, but those aspects of the subject which are directly related to the present experiments can be briefly summarized. It is sufficient to consider an equilibrium involving two components, whose mol fractions in the liquid phase are N_1, N_2, and in a solid phase in equilibrium with it $[N_1]$, $[N_2]$.

At equilibrium, the partial molal free energy of either component must be the same in the solid and liquid phase, e.g.

$$\bar{G}_1 = [\bar{G}_1] \quad \text{and} \quad \bar{G}_2 = [\bar{G}_2].$$

The simplest case arises when one of the components is practically insoluble in the solid phase, so that $[N_2] = 0$. Using standard results, $\bar{G} = \bar{H} - T\bar{S}$. When the concentration of the second component is small in the liquid phase, it is plausible to write for this phase

$$\bar{H}_1 = H_1^0, \tag{2}$$

where H_1^0 is the molar heat content of the pure liquid, and also

$$\bar{S}_1 = S_1^0 - R \log_e N_1, \tag{3}$$

where S_1^0 is the molar entropy. With these assumptions, the only effect of the component 2 in the liquid phase is to increase the partial molal entropy of the first component by the 'mixture' term $-R \log_e N_1$.

Neglecting changes in entropy and heat content due to changes of temperature, equilibrium will be established between (pure) solid and melt at a temperature T such that $\bar{G}_1 = \bar{H}_1 - T\bar{S}_1 = [\bar{H}_1] - T[\bar{S}_1]$.

Substituting from the above equations and transposing

$$T(S_1^0 - R\log_e N_1 - [\bar{S}_1]) = H_1^0 - [\bar{H}_1]. \tag{4}$$

For the pure solid $\quad [\bar{H}_1] = [H_1^0] \quad$ and $\quad [\bar{S}_1] = [S_1^0]$. (5)

Also $\quad \Delta H_f = H_1^0 - [H_1^0] \quad$ and $\quad \Delta S_f = \Delta H_f/T_M = S_1^0 - [S_1^0]$, (6)

where T_M is the melting-point of the pure solid in the absence of the second component.

Substituting in equation (4) and rearranging, it is seen that the above assumptions lead to the Hildebrand solubility equation

$$\Delta H_f(1/T_M - 1/T) = R\log_e N_1. \tag{7}$$

The experiments recorded in figures 9–10 show that when the pairs of components do not pack easily into a crystal lattice, so that solid solution is unlikely, this equation is obeyed in considerable detail, over the whole range of mixtures of long-chain compounds.

When the components can pack with ease into a crystal lattice, solid solution occurs. There is no reason to modify equations (2) and (3) for the liquid phase, but the equations (5) have to be modified, in view of the fact that $[N_2]$ is no longer negligible.

The most obvious assumption is to write for the solid lattice, by analogy with the liquid, whose structure is not so very different,

$$[\bar{H}_1] = [H_1^0] \quad \text{and} \quad [\bar{S}_1] = [S_1^0] - R\log_e[N_1], \tag{8}$$

and the solubility equation now becomes

$$\Delta H_f(1/T_M - 1/T) = R\log_e N_1/[N_1]. \tag{9}$$

When $N_2 = 1 - N_1$ and $[N_2] = 1 - [N_1]$ are both small, $R\log_e N_1/[N_1]$ is approximately equal to

$$R([N_2] - N_2) = RN_2([N_2]/N_2 - 1),$$

so that equation (9) corresponds with equation XX (19) given by Lewis and Randall (1923).

A comparison of this equation with experiment shows, however, that the assumptions from which equation (9) is derived cannot apply when the extent of solid solution is large. For example, in mixtures of *n*.hexadecane with *n*.hexadecene-1 (Langedijk and Brezesinska Smithuysen 1938),

$N_1/[N_1] \sim 0{\cdot}98$ for hexadecane when $N_1 \sim 0{\cdot}8$, so that there would be practically no depression of freezing-point according to equation (9). In actual fact, the freezing-point in ° K is found to obey very approximately a linear equation of the form

$$T = N_1 T_{M_1} + N_2 T_{M_2} = 290{\cdot}9 - 13{\cdot}8 N_2.$$

Similar equations are found to apply to other pairs of components with closely similar molecules, such as polymethylene hydrocarbons (Piper, Chibnall, Hopkins, Pollard, Smith and Williams 1931) and molecules built up from isotopic atoms (e.g. Kruis and Clusius 1938).

The interest of such mixtures for theories relating melting with crystal structure is that the mixture term $-R\log_e[N_1]$ does not change sufficiently on melting to have more than a subordinate effect on the equilibrium temperature. It has to be assumed that in composite crystal lattices the heat content $[H] = [N_1][\bar{H}_1] + [N_2][\bar{H}_2]$, and especially the entropy of the solid $[S] = [N_1][\bar{S}_1] + [N_2][\bar{S}_2]$, show a dependence on structure which is not allowed for in equations (8). To a first approximation, the entropy of mixing can be neglected altogether, and a plot of freezing-point against composition of the melt can be regarded as giving direct information on the relation between melting and structure. Various uses of this result are illustrated by the experiments described below.

Rotational factors in the melting of polymethylene lattices. A convenient type of lattice for investigating the relation between structural flaws and melting is that of polymethylene chains with a single CO group. The crystal structure of these compounds is frequently simpler than that of the corresponding paraffins, and can briefly be described by stating that the polymethylene chains lie parallel in the crystal, normal to the planes separating successive layers, and that the dipoles form sheets extending throughout the crystal. Owing to the fact that the polymethylene chains are not indefinitely long, the location of these sheets of dipoles with respect to the ends of the chain has a characteristic effect on the freezing-point (II), but if extreme cases are excluded their introduction into the lattice of heptadecane raises the freezing-point from about 295 to 314° K, and increases the entropy change on fusion from about 34·0 to 40·4 cal./mole/deg. (A similar comparison between $C_{18}H_{36}O$ and $C_{18}H_{38}$ is hampered by the different crystal structure of pure octadecane.)

When these layers of CO dipoles completely replace the corresponding CH_2 groups, they contribute an additional energy of about 5·3 kcal./mole or $3{\cdot}7 \times 10^{-13}$ erg per dipole. A rough picture of melting in this lattice would be that as the temperature is raised, at first a few dipoles begin to rotate, and that, owing to the resultant weakening of the co-operative field, the number rotating rapidly increases. The average attraction energy of freely rotating dipoles becomes comparable with that of CH_2 groups, so that when the dipoles are freed by rotation the lattice flaws correspond with those of heptadecane, and the crystal melts.

With this picture, the effect of temperature on the ketone lattice could be imitated by packing paraffin molecules of the same chain length into the crystal. These paraffin molecules would correspond with the rotating molecules in the lattice of the pure ketone The first few would have little effect on the melting-point, but as the co-operative field is weakened by the addition of more paraffin molecules an almost catastrophic increase of freedom must be reached at a critical concentration of the paraffin. On the other hand, if ketone molecules are packed into a paraffin lattice whose molecules are of the same chain length, they would not be expected to affect the melting-point appreciably until their concentration is high enough to lead to a co-operative field.

This picture of the melting of these composite lattices corresponds in some detail with the experimentally determined freezing-point curves of ketone/paraffin mixtures (figure 1). Throughout the series, experimental points are joined by broken lines. The Hildebrand freezing-point curves (continuous lines) which would apply if there were no solid solution have been included for comparison, and were calculated from equation (7), using previously published data (II, p. 331). The fact that solid solution takes place can be inferred from the freezing-points, which lie above the Hildebrand curves. On the basis of equation (9), $[N_2]/N_2 \sim 0{\cdot}75$ for small concentrations of ketone in paraffin, and $[N_1]/N_1 \sim 0{\cdot}2$ for small concentrations of paraffin in ketone. These are probably lower limits, but the amount of material was insufficient to estimate the ratios by an independent method.

From figure 1, it will be seen that for the ketones $C_{17}H_{34}O$ ($n = 3, 4, 5, 6$) the co-operative field appears to become significant at a concentration of about 30 dipoles %, whereas for $C_{17}H_{34}O$ ($n = 8$) the critical concentration is about 15 %. Both these estimates are necessarily very rough. The method of obtaining these curves is given below.

Defect flaws in the melting of polymethylene lattices. The simplest way of regarding a composite lattice is to start with pure crystals of one component, say $n.C_{17}H_{36}$, and to imagine what happens when a number of molecules are removed from the crystal, and are replaced by molecules of closely similar structure. Just as in the examples discussed for spherically symmetrical molecules (Wagner and Schottky 1930), two main types of flaw will result in the composite lattice. When the added molecule contains more methylene groups than the number removed, or when it contains side chains such as $(CH_3)_2CH.(CH_2)_{13}CH_3$, it will only be possible to pack it into the lattice of $n.C_{17}H_{36}$ by using the interstitial space, producing a considerable bulge in the regularity of arrangement. The energy associated with such 'interstitial' defects must be fairly high. Both theory and the experiments recorded below show that this type of composite lattice is difficult to realize.

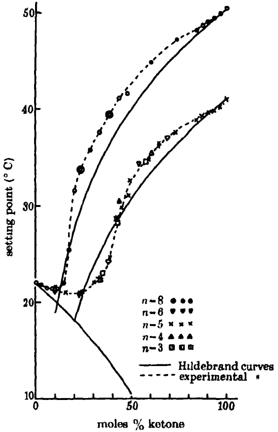


FIGURE 1. Freezing-point curves of composite lattices, ketone + paraffin. Upper curves. $C_{17}H_{36}+C_{17}H_{34}O$ ($n=8$). Lower curves: $C_{17}H_{36}+C_{17}H_{34}O$ ($n=3$, 4, 5 and 6).

On the other hand, when the added molecule contains fewer methylene groups than $C_{17}H_{36}$ it can be packed into the crystal lattice leaving holes, whose energy is less than $1{\cdot}1\times10^{-13}$ erg per missing group. A wide range of structures with such holes can readily be prepared (e.g. Piper *et al.* 1931, Langedijk and Brezesinska Smithuysen 1938). It should be noted, however, that a paraffin such as $C_{14}H_{30}$ could be moved anywhere along a vacant space in the lattice of $C_{17}H_{36}$, since the energy difference for three missing CH_2 groups will not vary greatly with the way the holes are distributed at the ends of the foreign molecule $C_{14}H_{30}$.

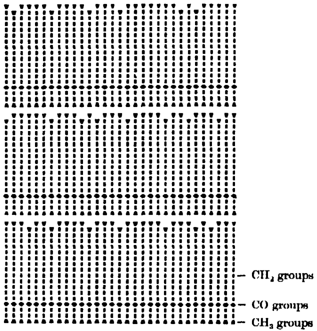


FIGURE 2. Composite lattice with random holes obtained by packing $C_{17}H_{34}O$ ($n = 3$) in $C_{18}H_{36}O$ ($n = 3$).

In order to obtain composite lattices with a more definite arrangement of defects, it was decided to 'anchor' the foreign molecules by working with mixtures of long chain ketones. The probability that the foreign ketone will have its CO group in the dipole sheet of the parent lattice can be estimated from the association energy to be about $e^{5300/RT} \sim 5000:1$. On this basis, various classes of defect lattices can be formed.

Defect lattices with one hole per foreign molecule. As is illustrated in figure 2, composite lattices of this type can be realized by packing together any one of the ketones $C_{18}H_{36}O$ ($n = 2, 3, 4, 5, 6, 7$ or 8) with the corresponding ketone of the series $C_{17}H_{34}O$, for example by combining $(C_3H_7)COC_{14}H_{29}$ with $(C_3H_7)COC_{13}H_{27}$, and so on. The methyl ketones have a different structure and cannot be used in the same way.

The experimental results for the freezing-point curves are given in figures 3 and 4. It will be noted from figure 3 that when the solid separating corresponds with $C_{17}H_{34}O$, the experimental points agree closely with the

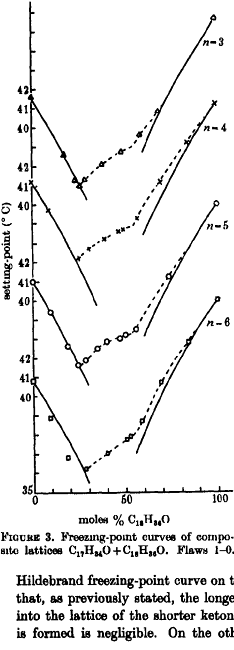


FIGURE 3. Freezing-point curves of composite lattices $C_{17}H_{34}O + C_{18}H_{36}O$. Flaws 1–0.

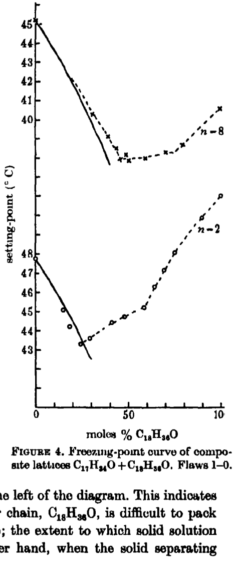


FIGURE 4. Freezing-point curve of composite lattices $C_{17}H_{34}O + C_{18}H_{36}O$. Flaws 1–0.

Hildebrand freezing-point curve on the left of the diagram. This indicates that, as previously stated, the longer chain, $C_{18}H_{36}O$, is difficult to pack into the lattice of the shorter ketone; the extent to which solid solution is formed is negligible. On the other hand, when the solid separating corresponds with $C_{18}H_{36}O$ (on the right of the diagram), the freezing-point curves lie above the Hildebrand curves. This indicates that considerable solid solution of the shorter chain can take place in the crystal of the longer chain, thus forming a lattice with holes.

In addition to this defect lattice, in which the holes are presumably distributed at random, it will be seen that a series of 'compounds', i.e. ordered solid solutions, are formed. Although compounds with non-congruent melting-points or with a melting-point maximum have been reported for a number of other long-chain molecules (*Annual Reports Chem. Soc.* 1939), in previous examples the associating group has always been at the end of the chain, so that it would be possible to pack the molecules in layers, without leaving any defects in the composite lattice. The present series is of particular interest in that the associating group is in the middle of the chain. For reasons discussed below, it has not yet proved possible to obtain X-ray spectrograms of these composite lattices, but it seems likely from energetic considerations that the molecules are packed so as to give an *orderly* arrangement of defects, in which the energy per hole is reduced by co-operative action.

The range of stability of this ordered lattice is of interest When the percentage of longer molecules is increased from 50 % to no more than 55 %, it becomes less stable than the (disordered) solid solution of $C_{17}H_{34}O$ in $C_{18}H_{36}O$, whereas the percentage of shorter molecules can be raised from 50 to 75 % before the composite lattice becomes less stable with respect to the melt, than (pure) $C_{17}H_{34}O$. This again emphasizes that it is comparatively easy to produce defect flaws in polymethylene lattices, but not to pack interstitial groups.

A further point which is suggested, but not completely proved, by the present data, is as follows Under comparable conditions, the entropy change on melting the ordered solid solution must be greater than that on melting the random solid solution of the same composition. Since the temperature of melting is given by $T_M = \Delta H_f/\Delta S_f$, the heat of fusion of the ordered lattice in the neighbourhood of $[N_1] = 0{\cdot}5$ must be larger than that of the random composite lattice, i.e. there is a co-operative action between lattice flaws which lowers their energy below that of a random arrangement. This conclusion is so important that further experimental evidence seems desirable. At present, the amounts of material available have been insufficient for direct determinations of the heat of fusion; information on the structure of the compounds might also be obtained by X-ray methods, but it has not yet been possible to obtain exposures in a sufficiently short time (see below).

Defect lattices with two or more holes per foreign molecule. Assuming that the CO groups tend to remain anchored in a composite ketone lattice, a number of different packing combinations can be realized which have more than one hole per foreign molecule. For example, two holes can be distributed with respect to the ends of the chain either 1–1 or 2–0.

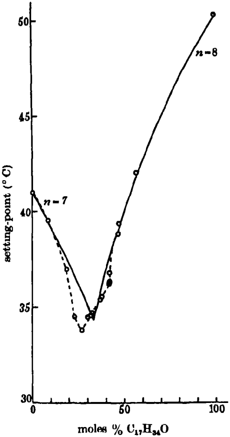


FIGURE 5 Freezing-point curve of composite lattice $C_{15}H_{30}O$ ($n=7$) + $C_{17}H_{34}O$ ($n=8$). Flaws 1–1.

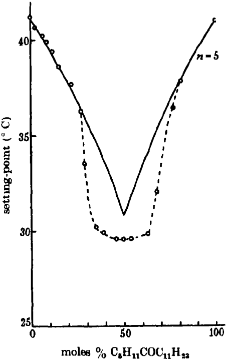


FIGURE 6. Freezing-point curve of composite lattice $C_{17}H_{34}O$ ($n=4$) + $C_{17}H_{34}O$ ($n=5$) Flaws 1–1.

The arrangement 1–1 has been investigated by mixing molecules of different chain length, $C_7H_{15}COC_7H_{15} + C_8H_{17}COC_8H_{17}$ (figure 5). It can also be realized by mixing ketones of the same chain length. Only the pair $C_{17}H_{34}O$ ($n=4+n=5$) is shown in figure 6, but other pairs have been investigated and show corresponding behaviour.

In both combinations, the normal freezing-point curve closely follows the Hildebrand equation. This indicates that any composite lattices which might separate have a lower freezing-point than the pure crystals throughout the whole range of compositions of the melt. Such a result might have been expected from the larger energy required to produce two holes in the lattice per foreign molecule.

In all three types of mixture it has, however, been possible to study the freezing-point curves of (metastable) defect lattices, as indicated by the broken curves. By analogy with the preceding curves, the side branches in figures 5 and 6 would correspond with the freezing-points of random solid solutions, and the lower cross branch the freezing-point of an ordered composite lattice with two defects per foreign molecule. The fact that the side branches of these metastable curves are depressed below the Hildebrand curves for the pure ketones indicates that the metastable lattices separating have a lower heat of fusion.

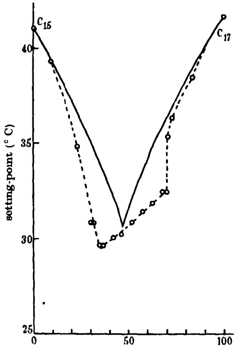


FIGURE 7. Freezing-point curve of composite lattice $C_{15}H_{30}O$ ($n = 7$) + $C_{17}H_{34}O$ ($n = 7$). Flaws 2–0.

The arrangement 2–0 has been realized by mixing the molecules $C_7H_{15}COC_7H_{15}$ and $C_7H_{15}COC_9H_{19}$ (figure 7), and again gives evidence of the separation of (metastable) defect lattices, including an ordered solid solution whose stability is greater than that of the random solid solutions over a considerable range of compositions. Details of the occurrence of these metastable phases are given below, but it may be stated here that their limited life before recrystallization sets in has so far precluded the investigation of their structure by X-rays

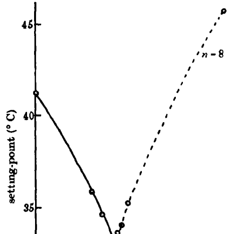


FIGURE 8

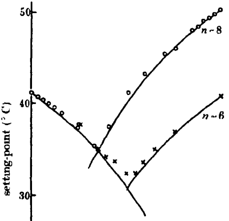


FIGURE 9

FIGURE 8. Freezing-point curve $C_{17}H_{34}O$ ($n = 4$) + $C_{18}H_{36}O$ ($n = 8$) (showing prolongation of solubility curve beyond the eutectic).

FIGURE 9. Freezing-point curves of mixed ketones. Upper curve: $C_{17}H_{34}O$ ($n = 4$) + $C_{17}H_{34}O$ ($n = 8$). Lower curve: $C_{17}H_{34}O$ ($n = 4$) + $C_{17}H_{34}O$ ($n = 6$).

The investigation of combinations of ketones which would pack with more than two holes per foreign molecule showed that the melting-points of such composite lattices are so much lower than the Hildebrand curves, that they are difficult to realize experimentally. The prolongation of the solubility curve of $C_8H_{16}COC_8H_{18}$ beyond the eutectic in figure 8 is noteworthy. This crystallization of a metastable phase seems fairly common in polymethylene lattices (cf. other examples in this paper, and Keffley and Maiden 1936). It provides further illustration of the rule of successive states (cf. (III)), and its frequent occurrence is probably connected with the large entropy change on freezing these melts. Figure 8 also shows how closely the combination 4–3 obeys the Hildebrand equation, and figure 9 illustrates the results for the combinations 2–2 and 4–4.

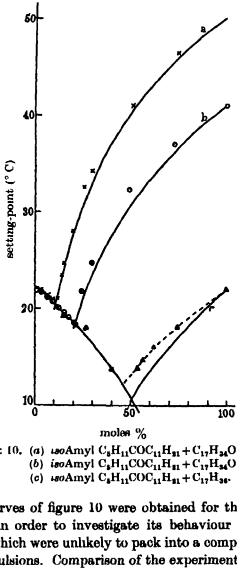


FIGURE 10. (*a*) *iso*Amyl $C_5H_{11}COC_{11}H_{21}+C_{17}H_{34}O$ ($n=8$).
(*b*) *iso*Amyl $C_5H_{11}COC_{11}H_{21}+C_{17}H_{34}O$ ($n=5$).
(*c*) *iso*Amyl $C_5H_{11}COC_{11}H_{21}+C_{17}H_{36}$.

Finally, the curves of figure 10 were obtained for the *iso*amyl ketone, $C_5H_{11}COC_{11}H_{21}$, in order to investigate its behaviour when mixed with other molecules which were unlikely to pack into a composite lattice owing to interstitial repulsions. Comparison of the experimental points with the Hildebrand curves shows that solid solution is negligible between this ketone and $C_{17}H_{34}O$ ($n=8$). The *iso*ketone dissolves to a small extent in the lattice of $C_{17}H_{34}O$ ($n=5$), but not vice versa. It also dissolves to a measurable extent in the lattice of $C_{17}H_{36}$, but not vice versa. All these examples can be summed up by the statement that interstitial flaws are difficult to produce in these lattices.

CRYOSCOPIC DATA

Owing to the necessity for economy in the use of materials, the majority of the freezing-point curves were determined by a previously described method (II, p. 330) New cryoscopic data are recorded in table 3. These data were obtained for compounds prepared and purified by tested methods (II; Oldham and Ubbelohde 1939*a*), though the purity of the products could not be verified as fully as in the series $C_{17}H_{34}O$. This might mean that the above values are too high, owing to the spurious rise in cryoscopic heats of fusion when homologues favour solid solution (cf. II, p. 332). A comparison with other data is nevertheless of interest for theories of the relation between melting and crystal structure Thus the introduction of one *iso*methyl group in $C_{17}H_{34}O$ lowers the entropy of fusion by about 6·6 units, and the heat of fusion by about 2×10^{-13} erg per molecule. Again, the ketone $C_{18}H_{36}O$ ($n = 5$) has an entropy of fusion some 5·2 units greater than $C_{17}H_{34}O$ ($n = 5$). This increase is considerably greater than the difference per CH_2 group for homologous 'vertical' paraffins (Garner's formula gives about 2 entropy units per CH_2 group). Finally, the ketone $C_7H_{14}COC_7H_{14}$ has an entropy of fusion about 5·5 units lower than the ketone $C_8H_{16}COC_8H_{16}$.

TABLE 3

Substance	Solute	Entropy of fusion (cal./mol./deg.)	Heat of fusion (kcal./mole)	T_M (° C)
$C_{18}H_{36}O$ ($n = 5$)	$(C_6H_5CH_2)_2$	45·7	14·6	45·0
$C_{17}H_{34}O$ ($n = 5$, *iso*amyl)	Other ketones	33·8	10·0	22·4
$C_{15}H_{30}O$ ($n = 7$)	$(C_6H_5CH_2)_2$	36·8	11·6	41·0

Life of metastable phases As a preliminary to X-ray investigations, it was necessary to determine the stability of the composite lattices whose formation had been detected from the freezing-point curves. Although the X-ray investigations have been deferred, the nature of the results may be briefly indicated.

Using the technique described in (II), a sealed tube containing a melt of the appropriate composition was quickly cooled to produce a small quantity of solid. In the case of the metastable phases this solid had a transparent appearance quite different from that of the pure ketones. The persistence of this solid was tested by plunging the tube into a bath maintained at various temperatures. The freezing-point curves mark upper temperature limits above which the phase would not persist. Some practice is required to obtain good results, which depend in part on the slow growth of solid from liquid near the freezing-point. In the case of metastable lattices, the solid was found to melt freely above the temperatures indicated by the broken curves, provided the tube was not kept too long at a lower temperature before testing the persistence. If a critical period was exceeded, then on plunging the tube into the testing bath only partial melting of the solid took place and was immediately followed by hazing and resolidification of the contents When this happened, melting could not be completed till the testing bath was raised above the Hildebrand temperature for the melt. The critical life of the metastable lattice limits the experiments which can be performed to elucidate its structure. Typical results are reproduced in table 4

TABLE 4. RECRYSTALLIZATION OF (ORDERED) DEFECT LATTICES

Composition $C_{17}H_{34}O$ $(n=4+n=5)$	Freezing-point °C	Testing bath °C	Cooling °C	Life min
$N_1 = 0·40$	29·9	+1·1	−0·9	5
0·47	29·7	+1·0	−0·7	70
			−1·2	38
			−1·8	15
			−3·0	6
0·50	29·6	+1·0	−1·2	30
0·55	29·7	+1·4	−0·7	17
0·63	29·8	+1·2	−0·8	1

In this table, the third column indicates the temperature of the testing bath, above the observed freezing-point, and the fourth column the temperature at which the tube was maintained in order to measure the life of the phase. These and other experiments showed that this life decreases rapidly with increasing cooling below the freezing-point, and also as the composition of the solid departs more and more from the 1 : 1 ratio

DISCUSSION

The main conclusions from the experiments are·

(1) In cryoscopic determinations of the latent heat of fusion near the melting-point of the pure compound, a drift of about 9 % is observed for octadecane, and a drift not exceeding 3·5 % for dioctyl ketone.

(2) Freezing-point curves for mixtures of a ketone with a paraffin of the same chain length indicate that whereas ketone molecules can pack freely into the paraffin lattice, without much affecting the transition temperature from solid to liquid, the paraffin molecules at first pack with some difficulty into the ketone lattice. Beyond a limiting concentration of paraffin there is a 'catastrophic' breakdown, comparable with that which occurs on melting a ketone.

(3) Freezing-point curves for pairs of closely related ketones indicate that both ordered and random defect lattices are much more readily formed than solid solutions with interstitial groups. When the number of holes does not exceed two per foreign molecule, or 12 % of all the methylene groups, ordered solid solutions are formed whose maximum melting-point is not very different from that of the next lower ketone. When these ordered composite lattices are metastable, their life diminishes rapidly on cooling below the freezing-point, or on varying the composition from the 1 : 1 ratio.

From these conclusions it is possible to make the following inferences:

(1) The cryoscopic data support the previously made suggestion that for octadecane lattice flaws must make an appreciable contribution to the lattice energy in the neighbourhood of the melting-point. This implies that the whole of the structural entropy change on passing from solid to liquid does not take place at one temperature. The present experiments do not exclude the possibility of an isothermal entropy increase at a true melting 'point', in addition to the gradual increase up to the melting-point; in fact, they are most readily interpreted on this basis. Furthermore, the cryoscopic data on octadecane do not decide what type of crystal flaw is chiefly responsible for melting.

(2) The cryoscopic data on ketones lead to a more definite picture of lattice flaws of importance for melting. Composite lattices of a ketone with a paraffin of the same chain length behave in a way which indicates that a rotation of the CO group favours melting. Composite lattices of pairs of ketones of closely similar structure behave in a way which indicates that holes are readily formed in polymethylene lattices. The waxiness of these compounds may be connected with this property.

On this basis, a paraffin lattice such as octadecane might begin to increase its structural entropy by forming lattice holes Isolated holes could only be formed in sequences parallel to the long axis of the molecules, but the behaviour of ordered defect lattices strongly suggests that a crystal can lower its lattice energy by grouping its thermal flaws into a co-operative system. A single crystal traversed by such a network of co-operative flaws would not differ appreciably from the crystal mosaic familiar to X-ray spectroscopists. An extreme statement of a micelle theory of melting would compare the phenomenon with a sol-gel transition in which the micelles were crystalline, and in which the intermicellar molecules could be built into the crystalline pattern at lower temperatures. In certain organic crystals the energy required to produce flaws may become so small that the physical properties are visibly affected (cf. perhaps the short-lived transitional phases in $C_{32}H_{66}$, Seyer and Morris 1939) and certain infra-red absorption spectra near the melting-point (Taschek and Williams 1939).

Melting and the mechanical properties of solids

The view that thermal flaws of a co-operative kind are a necessary feature of crystal lattices in thermodynamic equilibrium (cf. I) receives additional support from the examples of co-operative defects described in this paper. Owing to the relatively slow attainment of equilibrium in co-operative systems (III) it is difficult to make sure that any crystal as actually observed has reached a state with the true minimum of free energy. Nevertheless, it simplifies the problem to discuss what would be observed if it were possible to realize the crystal lattice with minimum free energy, right down to 0° K Crystal flaws can then be classified according to their contribution to energy and entropy terms in the expression for the free energy.

Contributions to the energy at 0° K. The possibility cannot be neglected that at 0° K a mosaic or lineage structure may have a smaller heat content than an infinite perfect lattice. Various suggestions do not seem to be admissible (cf. Buerger 1934), but effects connected with the zero-point energy may still be of importance. For example, a formal analogy between the reflexion of electrons and X-rays by ordered arrangements of atoms (cf. Mott and Jones 1936) may extend to their behaviour with regard to a crystal mosaic For X-rays, the result of breaking up a perfect crystal into a mosaic (Darwin 1914, 1922; Compton and Allison 1935) is to increase the range of wave-lengths which cannot penetrate appreciably into the lattice, owing to Bragg reflexions A similar effect for the electron waves within a crystal would widen the range of forbidden energies in the case of a mosaic. For crystals with comparatively wide energy bands, such as metals, the modification in lattice energy on forming a mosaic, due to this change in the bands, might be sufficiently important to make it the most stable crystal structure at 0° K. On this basis it might be possible to account for the plasticity of 'single crystals' of metals at 0° K, and to explain such phenomena as the limiting size of crystallites on plastic deformation (Wood 1939), and the fact that the tighter binding of a cold worked metal is in parallel with the increased scattering of electron waves (i.e. the electrical resistance) at 0° K.

The reason for including a discussion of possible crystal flaws at 0° K in dealing with melting is that whereas raising the temperature increases the proportion of thermal flaws, as mentioned in the next section, an approach towards 0° K may increase the flaws due to zero-point energy effects, so that certain crystals would never exhibit perfect lattices when in equilibrium (cf. perhaps the case of Na, Dawton 1937).

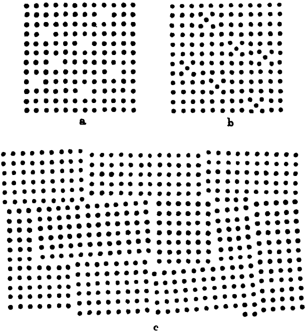


FIGURE 11. Thermal origin of lattice flaws. (*a*) Isolated holes in a lattice (*b*) Interstitial atoms packed into the lattice (*c*) Co-operative flaws combining both features, to form a crystal mosaic.

Contributions to the energy and entropy at higher temperatures. In addition to possible energy effects at 0° K, certain lattice flaws must exist in equilibrium at higher temperatures, owing to the increased entropy of the crystal when they are present. Figure 11*a* illustrates the formation of isolated holes in a lattice, and figure 11*b* the packing of atoms in the interstices (for these lattice flaws cf. Wagner and Schottky 1930).

Figure 11*c* illustrates how the co-operative flaws in a mosaic lattice combine the characteristics of both types of isolated flaw. Since the mutual potential energy of atoms is governed by an expression in which the repulsion energy changes much more rapidly with distance than the attraction energy, the mosaic has been drawn so as to make a decrease of interatomic distance comparatively rare. On this view, the average density of a mosaic should be smaller than that of perfect regions within it. It will be clear from the diagram that the flaw energy per atom can be made much smaller in a co-operative flaw than in isolated flaws. On the other hand, a co-operative action between lattice defects must lessen the number of ways in which they can be selected in the crystal, and must thus lessen their contribution to the entropy. Qualitatively, the result will be that although co-operative flaws can become important in the crystal at lower temperatures than isolated holes, they may not increase as rapidly with rising temperature.

Lattice flaws and melting. The conventional treatment of melting from the standpoint of the phase rule is most conveniently discussed in terms of the thermodynamic potential G at constant pressure. This quantity is greater for the liquid than for the solid phase below the melting-point. The G, T curves for solid and liquid have no necessary connexion with one another, so that they cut at a sharp angle at the melting-point T_M. The entropy change $\Delta S = dG_S/dT - dG_L/dT$ and the volume change $\Delta V = -dG_S/dP + dG_L/dP$ occur isothermally at T_M, and a plot of any property such as the heat content against the temperature would show an abrupt change at the melting-point.

This conventional treatment cannot be regarded as experimentally verified. Even for quite simple lattices experiments do not show that melting is complete at a melting 'point'. For example, the volume change on melting specially pure mercury (Smits and Muller 1937) takes place over a range of 0·024°. This can no doubt be explained on the usual assumption of Raoult impurities, but might equally well be due to the fact that a sharp phase change is in reality a phase change 'of the second kind' taking place over a sufficiently narrow range of temperatures. The mathematical treatment of the two processes (Rutgers and Wouthuysen 1937) is so similar that the experimental foundation of the phase rule requires careful verification. It would be rash to attempt a reformulation without further experimental information as to what actually happens in the transition from solid to liquid, but some of the peculiarities of fluctuations in co-operative systems must be of importance (cf. III, and Ubbelohde 1937*b*, p. 300).

If co-operative lattice flaws are of importance for melting, as is suggested in this paper, it seems difficult to believe that the solid-liquid transition can be entirely free from hysteresis, though it might be difficult to detect. The subdivision of a crystal lattice into a micellar aggregate, by a network of such flaws, may perhaps be regarded as a more concrete formulation of 'heterophase fluctuations' (Frenkel 1939; Brody 1939).

Experimental details of preparation and purification

Octadecane. In order to have a check on the absence of homologues, this hydrocarbon was prepared in two independent ways·

(*a*) Octadecyl alcohol (Th. Schuchardt) was recrystallized from ethyl alcohol, and then from acetone. With each solvent, recrystallization was continued until the setting-point of the substance recovered, on evaporating the mother liquors, converged to within ± 0·1° of the setting-point of the crystals. Essentially, this method of testing for purity uses the fact that with a suitable solvent the partition coefficient of an impurity between mother liquor and crystals can be much greater than unity. The limiting value of the setting-point of the crystals (57·9 ± 0·1°) was reached one or two operations sooner than this convergence point, and is thus a less stringent test. The purified alcohol was then converted to iodide, which was recrystallized in a similar manner from alcohol and acetone (limiting freezing-point 33·0 ± 0·1°), and was then converted to hydrocarbon and purified as previously described (Oldham and Ubbelohde 1938). The final setting-point of a sample prepared in this way was determined by immersing a calibrated micro-Beckmann thermometer in the melt, and was 28·021 ± 0·005° (Smith 1933 gives melting-point 28·02°). The chief impurities in this sample would be traces of closely similar hydrocarbons, and of $C_{36}H_{74}$, which might form solid solutions

(*b*) In order to have a comparison sample without these impurities, octadecane was also prepared from nonyl iodide by a modified Grignard synthesis (Oldham and Ubbelohde 1938). Nonyl alcohol (Th. Schuchardt) was converted to the *p*. dinitrobenzoic ester by heating with the acid chloride at about 110° C in a large flask This ester was recrystallized from petroleum ether, from alcohol, and from glacial acetic acid, using the convergence test, till a limiting setting-point was reached of 48·0 ± 0·1°. The ester was then converted back to nonyl alcohol, and then to iodide. The object of this circuitous process was to eliminate chain lengths closely similar to octadecane at the half-way stage, where the difference in physical properties is more pronounced, and hence the purification more definite. The nonyl iodide was then 'doubled' by the Grignard method. The product after fractionating *in vacuo* was treated with Nordhausen acid, and recrystallized from alcohol and acetone, using the convergence test. It was then again treated with Nordhausen acid, followed by conc H_2SO_4 and water, in order to remove any traces of organic solvents which might be retained in the lattice. This precaution was adopted since it was observed that when a paraffin is recrystallized from CCl_4 the crystals give a positive test for chlorine for a considerable time, probably because the CCl_4 is retained in holes in the lattice. Finally, the octadecane was fractionated *in vacuo*. The final setting-point on immersing a calibrated micro-Beckmann thermometer was found to be 28·096 ± 0·005°. Only this sample was used in the cryoscopic experiments. Its homogeneity of chain length was confirmed by the fact that two further treatments with Nordhausen acid (to remove added solutes) followed by purification gave final products with the setting-points 28·095 ± 0·005 and 28·090 ± 0 005.

Dioctyl ketone. Owing to the comparatively large amounts required, this compound was prepared by distilling zinc pelargonate *in vacuo* at about 420° C. The distillate was dissolved in benzene, and washed with alkali and water. Esters were removed by boiling with alcoholic NaOH, and the product was fractionated *in vacuo* (178–180°, 5 mm.) The crude ketone obtained in this way was recrystallized from alcohol and acetone, using the convergence test. The final setting-point measured with a calibrated micro-Beckmann thermometer immersed in the melt was 50·230 ± 0·005°.

Other ketones were prepared by previously described methods, and were recrystallized from alcohol and acetone till the freezing-point of the crystals converged to within ± 0·1° C with that of the more soluble fractions.

Thanks are due to Miss I Woodward for her valuable assistance with the diagrams, to Professor J. Alexander, Professor J. Frenkel and Dr R. S. Jessup for various communications, and to the Managers of the Royal Institution for the facilities provided.

REFERENCES

Annual Reports Chem. Soc. 1939 **35**, 262

Born, M 1939 *J Chem. Phys.* **7**, 591.

Brody, E 1939 *J. Chem. Phys.* **7**, 538

Buerger, M J. 1934 *Z. Kristallogr.* **89**, 195.

Compton and Allison 1935 *X-rays in theory and experiment*, pp. 366 *seq.* Macmillan.

Darwin, C. 1914 *Phil. Mag* **27**, 325, 675.

— 1922 *Phil. Mag.* **43**, 800.

Dawton, R. 1937 *Proc. Phys. Soc.* **49**, 294.

Debye, P. 1939 *Z. Elektrochem.* **45**, 174.

Frank, F. C. 1939 *Proc. Roy. Soc.* A, **170**, 182.

Frenkel, J. 1939 *J. Chem. Phys.* **7**, 538.

Guggenheim, E. 1937 *Trans. Faraday Soc.* **33**, 151.

Hildebrand, J. H. and Sweeny, J 1939 *J. Phys. Chem.* **43**, 297.

Irany, E. 1939 *J. Amer Chem. Soc.* **61**, 1436.

Jost, W. 1933 *J. Chem. Phys.* **1**, 466.

Keffley, L. and Maidon, A. 1936 *J. Phys. Chem.* **40**, 905.

Kruis, A. and Clusius, K. 1938 *Z. Phys. Chem.* B, **38**, 156.

Langedijk, S. and Brezesinska Smithuysen, W. 1938 *Rec. Trav. Chim. Pays-Bas*, **57**, 1050

Lennard Jones, J. E. and Devonshire, A. 1939 *Proc. Roy Soc* A, **170**, 464.

Lewis and Randall 1923 *Thermodynamics*, p. 238. McGraw Hill Book Co.

Lucas, R. 1938 *C.R. Acad Sci., Paris*, **207**, 1408.

Mott, N. F. and Gurney, R. W 1939 *Trans. Faraday Soc.* **35**, 364.

Mott, N. F. and Jones, H. 1936 *The electron theory of metals*, pp. 63, 64. Oxford

Oldham, J. W H. and Ubbelohde, A. R. 1938 *J Chem Soc* 201.

— — 1939*a* *J. Chem. Soc* 201.

(II) — — 1939*b* *Trans Faraday Soc.* **35**, 328.

Piper, S. H , Chibnall, A. C., Hopkins, S. J., Pollard, A , Smith, J. A B. and Williams, E. F. 1931 *Biochem. J.* **25**, 2087.

Rutgers, A. and Wouthuysen, S 1937 *Physica*, **4**, 235, 515

Seyer, W F. and Morris, W. 1939 *J. Amer. Chem Soc.* **61**, 1114.

Smith, J. C 1932 *J. Chem. Soc* 741.

— 1933 *J. Chem. Soc.* 346.

Smits, A. and Muller, G. J. 1937 *Z. Phys. Chem* B, **36**, 288.

Taschek, R and Williams, D. 1939 *J. Chem Phys* **7**, 11.

(III) **Ubbelohde** 1937*a* *Trans. Faraday Soc.* **33**, 1203.

— 1937*b* *Proc. Roy. Soc* A, **159**, 295.

(I) — 1938 *Trans. Faraday Soc* **34**, 299.

— 1939 *Trans. Faraday Soc.* **35**, 843.

Wagner, C and Schottky, W. 1930 *Z. Phys Chem.* B, **11**, 163.

Wannier, G. 1939 *J. Chem. Phys.* **7**, 810.

Wood, W. A. 1939 *Proc. Roy Soc.* A, **172**, 231.

Isomorphous replacement and superlattice structures in the plagioclase felspars

By S. H. Chao and W. H. Taylor
Physics Department, College of Technology, Manchester

(*Communicated by W. L. Bragg, F.R S.—Received* 19 *April* 1940)

From X-ray oscillation and rotation photographs of a limited number of plagioclase felspars it seems probable that the structure type depends upon the chemical composition in the following way. From pure soda felspar to at least 22 % lime felspar there extends an isomorphous series with the *albite* structure, based on a unit cell with linear dimensions 8 × 13 × 7A approximately and containing four molecules of $NaAlSi_3O_8$ or the equivalent when (CaAl) replaces (NaSi). There is probably a similar range of isomorphism, near pure lime felspar $CaAl_2Si_2O_8$, with the *anorthite* structure which has axes 8 × 13 × (2 × 7)A, but is of course very similar to the albite structure The characteristic features of the photographs of the intermediate plagioclases can be explained, with the help of an optical model, in terms of a structure in which sheets with the albite structure and sheets with the anorthite structure alternate, but more experimental data are needed for the complete elucidation of the problem.

1. Introduction

The plagioclase felspars are generally regarded as a continuous series of isomorphous mixtures, in varying proportions, of the end-members albite ($NaAlSi_3O_8$) and anorthite ($CaAl_2Si_2O_8$). They are immediately distinguished from the orthoclase felspars by the fact that the cleavages (001) and (010), right-angled in the latter, are inclined at about 85° in the plagioclases. This difference between a typical plagioclase and a typical orthoclase is a consequence of the more complete collapse of the tetrahedron framework around the smaller sodium or calcium ions in the former, the main structural scheme being the same in all felspars (Machatschki 1928; Taylor, Darbyshire and Strunz 1934). Features of the X-ray data for different members of the plagioclase group led to the suggestion (Taylor *et al.* 1934) that these minerals may constitute not a single but a double isomorphous series. The present paper records observations which indicate that a more complex hypothesis is necessary and attempts to outline an explanation of the general features of the experimental data.

2. EXPERIMENTAL DATA

Table 1 summarizes the results of our experimental measurements of densities, extinction angles on (010), and axial lengths of the crystallographic unit cell. The figures quoted in the table must be interpreted in the light of the detailed description of the effects observed with each specimen.

TABLE 1. AXIAL LENGTHS OF PLAGIOCLASE FELSPARS

		Percentage composition				Axial lengths in A			
Specimen	Origin	Ab	An	Or	Extinction on (010)	*a*	*b*	*c*	Density
A, Albite	Auburn, Maine	91·0	5·2	0·8	19°	8·12	12·85	7·17	2·62
B, Oligoclase	Haddam, Conn	90·5	9·5	—	12°	(8·1)	12·84	7·16	2·63
C, Oligoclase	Kraegerö, Norway	77·2	21·8	1·0	8°	8·11	—	7·16	2·64
D, Andesine	Korea	~70	~30	—	−12°	—	—	(7·14)	2·66
E, Labradorite	Naine, Labrador	41·2	57·1	1·7	−20°	8·17	12·85	7·10	2·70
F, Labradorite	Russia	~35	~55	~10	−20°	(8·2)	12·86	7·11	2·69
G, Bytownite	Crystal Bay, Minn.	32·8	67·2	—	−31°	8·17	12·88	7·11	2·60
H, Bytownite	Crystal Bay, Minn	19·8	79·2	1·0	−31°	—	—	7·11	2·72
I, Anorthite	Mikyinna, Japan	0·3	99·1	0·6	−36°	8·17	12·88	7·09	2·76

The chemical compositions are quoted as percentages of albite $NaAlSi_3O_8$, anorthite $CaAl_2Si_2O_8$, and orthoclase $KAlSi_3O_8$

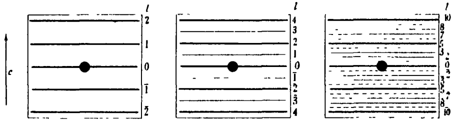
A blank space in the columns headed 'axial lengths' indicates that no measurement has been made, and a numerical value enclosed in brackets is known to be less reliable, or less accurate, because of some special experimental difficulty

We are indebted to Professor C. D. Jeffries of the State College, Pennsylvania, U.S.A., for all the materials except specimens D and F and for the information as to place of origin and chemical composition quoted in table 1. We also wish to thank Professor R. C. Emmons of the University of Wisconsin for his help in arranging the supply of specimens. We are grateful to Dr E. Spencer for specimen F and for the chemical analysis of this material, and to Professor S. Kôzu for specimen D. No chemical analysis is available for this Korean andesine, and the composition quoted in the table is estimated from the measured density and extinction angle on (010).

The densities were determined by flotation, and are probably accurate to ±0·01.

The unit cell dimensions were determined from X-ray rotation and oscillation photographs. When a small single crystal is mounted with a crystallographic axis parallel to the vertical rotation axis, the reflexions registered on a cylindrical film are arranged on a series of straight horizontal layer-lines: the separation of the layer-lines, measured in reciprocal lattice co-ordinates, is inversely proportional to the length of the crystallographic axis about which the crystal rotates (or oscillates) (see figure 1). For a few photographs molybdenum K_α radiation ($\lambda = 0{\cdot}710$A) was used, and for the rest copper K_α radiation ($\lambda = 1{\cdot}54$A).

The experimental materials appear to fall into three groups, which are described below. Before proceeding to discuss the features which differentiate one group from another, it is, however, necessary to emphasize that to a first approximation *all* the structures are the same, corresponding X-ray photographs being very similar in general appearance and differing only in detail.



albite $c=7$A anorthite $c=14$A labradorite $c=35$A

Figure 1. Diagrammatic representation of c-axis rotation photographs. Individual reflexions are not shown, the positions of the observed layer-lines (in reciprocal lattice co-ordinates) being represented by continuous lines. For simplicity, the length of the c-axis is taken to be 7A in albite, and a multiple of 7A in other plagioclases, the difference between the lengths of the albite true axis (7·16A) and anorthite pseudo-axis (7·10A) being neglected. Principal layer-lines (drawn in thick line) are in the same positions in the three photographs, but are numbered differently in accordance with the length of the *true* axis in each case. Single subsidiary layer-lines (in thin line) appear midway between principal layer-lines in anorthite, indicating an axis of length 2 × 7A. Pairs of subsidiary layer-lines appear between principal layer-lines in the labradorite specimen represented, and correspond in position to the 2nd and 3rd, 7th and 8th, etc., layer-lines for an axis of length 5 × 7A. Dotted lines indicate the positions of layer-lines not observed experimentally, or containing at most a few exceedingly feeble reflexions

(i) *Albite and isomorphous structures*

Specimens A, B, C represent an isomorphous series with the albite structure, based on a unit cell with axes $a=8{\cdot}11$A, $b=12{\cdot}85$A, $c=7{\cdot}16$A, and with interaxial angles which presumably vary slightly according to the composition of the material but are not determined in the present work. The unit cell contains *four* molecules of composition $NaAlSi_3O_8$ in the case of pure albite, and when lime felspar is present replacement of (NaSi) by (CaAl) takes place without serious disturbance of the structure. The limit of the range of isomorphism is not fixed with any accuracy: it extends as far as C with 22% lime felspar, but not as far as E with 57% lime felspar; the data for D ($\sim 30\%$ lime felspar) are not sufficiently complete to determine whether this material falls within the isomorphous series.

Peculiarities of the photographs which appear to be of minor importance, and which await further investigation, are as follows: in a-axis photographs there is a tendency (apparently systematic) for some reflexions to be very slightly displaced above or below the layer-line on which they should be centred, which results in layer-lines of rather uneven appearance; and in *b*-axis photographs we have observed very weak satellites, which accompany the stronger reflexions, appear in the same positions in photographs from different specimens, and may perhaps be due to partial twinning.

(ii) *Anorthite and isomorphous structures*

Specimens I and H probably represent an isomorphous series with the anorthite structure, differing slightly from the albite structure, and based on a unit cell with axes $a = 8{\cdot}17$A, $b = 12{\cdot}9$A, $c = 2 \times 7{\cdot}10$A. (The inter-axial angles are not determined by our methods of measurement.) This unit cell contains *eight* molecules of composition $CaAl_2Si_2O_8$ in the case of pure anorthite, and replacement of (CaAl) by (NaSi) is tolerated (within the range of isomorphism) without serious disturbance of the structure.

The difference between the albite and anorthite structures is revealed first by small differences between corresponding axial lengths, and secondly by the doubled cell in anorthite. Thus we have 8·11 and 8·17A, 12·85 and 12·9A, 7·16 and $2 \times 7{\cdot}10$A, for *abc* axes in albite and anorthite structures respectively. These may be compared with the data published by Taylor *et al.* (1934) for quite different materials, viz. 8·14 and 8·21A, 12·86 and 12·95A, 7·17 and $2 \times 7{\cdot}08$A, for *abc* axes in albite and anorthite—figures which are slightly different in absolute magnitude but show precisely the same relationships. The periodicity of approximately 7A along the direction of the crystallographic *c*-axis, which is the true *c*-axial length in the albite structure, is only a pseudo-axis in the anorthite structure; for a single set of subsidiary layer-lines appears midway between the principal layer-lines, indicating that the true *c*-axis is *twice* the pseudo-axis (see figure 1). The closeness with which the pseudo- repeat of 7·10A approaches an exact repeat is indicated by the relatively low intensities of the reflexions lying on the subsidiary layer-lines. These differences between anorthite and the isomorphous series of structures of the albite type appear to be firmly established.

The data at present available are not, however, sufficient to enable us to state with equal certainty that a whole series of isomorphous structures of the anorthite type exists up to 20% soda felspar (specimen H). For although in *c*-axis photographs of this material there are well-marked subsidiary layer-lines midway between the principal layer-lines (exactly as in pure anorthite, specimen I), in addition a few much weaker reflexions, apparently midway between principal and subsidiary layer-lines, can be detected. Only if these exceedingly feeble reflexions are neglected—their real existence is in fact rather uncertain—can we suppose the structure of specimen H to be effectively identical with that of pure anorthite, but if this is accepted, the limit of the range of isomorphism is fixed between 20% soda felspar (specimen H) and 33% soda felspar (specimen G); for the structure of the latter is definitely different from that of anorthite (see below).

Further specimens must be examined to confirm the existence of an isomorphous series with the anorthite structure, and to determine the limit of the range of isomorphism.

(iii) *Intermediate plagioclases*

Specimens E, F, G represent a series of structures based on a pseudo-cell with the *anorthite* axial lengths, $a = 8{\cdot}17$A, $b = 12{\cdot}9$A, $c = 7{\cdot}10$A, the series extends on the lime-rich side to 67% lime felspar (G) but not to 79% lime felspar (H); the boundary on the soda-rich side lies within the wide range between 41% soda felspar (E) and 77% soda felspar (C)—we have been unable to decide with any certainty whether D, with $\sim 70\%$ soda felspar, belongs to this group or to the albite structures.

The following features appear to be characteristic of the photographs of members of this group of intermediate plagioclases. *First*, in *a*-axis and *b*-axis photographs there are very weak subsidiary reflexions which may be interpreted as indicating the existence of components of the crystal structure with (subsidiary) axes a_2, a_2' and b_2, b_2', inclined at small angles ($\sim 1°$) to the principal axes a_1 and b_1 respectively, and of the same lengths as these axes within the accuracy of our measurements. These subsidiary reflexions are much less intense in G than in E and F. *Secondly*, in *c*-axis photographs there are *pairs* of subsidiary layer-lines lying symmetrically on either side of a position midway between principal layer-lines. With specimen D a few faint markings may represent subsidiary reflexions, but they are so indefinite that they cannot usefully be discussed. In photographs of specimens E and F well-marked subsidiary layer-lines appear* in positions which correspond at least to a first approximation to the 2nd and 3rd, 7th and 8th, etc., layer-lines of a *true* c-axis of length $5 \times 7{\cdot}1$ A, with the principal layer-lines as zero, 5th, 10th, etc. A few faint markings may represent reflexions lying on the 1st and 4th, 6th and 9th, etc., layer-lines, but these are exceedingly feeble and may not represent reflexions at all (see figure 1). For specimen G there are subsidiary layer-lines in positions which correspond (within the accuracy of our measurements) to the 3rd and 4th, 10th and 11th, etc., layer-lines for a *true* c-axis $7 \times 7{\cdot}1$ A, with principal layer-lines as the zero, 7th, 14th, etc., layer-lines of the true axis, and with a few doubtful reflexions (of exceedingly low intensity) on the remaining layer-lines. The subsidiary reflexions of G are on the whole definitely more diffuse, and weaker in relation to the principal reflexions, than those of specimens E and F (cf. a- and b-axis photographs, above).

---

* Taylor *et al* (1934, pp. 478, 479) report a specimen of labradorite from Madagascar as having a true cell identical with that of anorthite, though the individual subsidiary reflexions are different in the two crystals. The density is 2·698, and the extinction on (010) is 20°, values which suggest a material similar to E and F, and we find on re-examination of the material that the *c*-axis photographs are in fact similar, with subsidiary reflexions lying on a pair of layer-lines of about the same separation as in E and F. The earlier workers, using molybdenum radiation and flat plates, failed to notice the slight separation of the layer-lines. The length of the pseudo-axis is 7·10A, in agreement with the measurements of E and F.

We do not know whether the separation of the two members of a pair of subsidiary layer-lines varies continuously with composition, or in a succession of discontinuous steps each of which covers a limited but finite range of compositions. The layer-lines are more widely separated for the soda-rich specimens E and F ('true' c-axis $5 \times 7{\cdot}1$ A) than for G ('true' c-axis $7 \times 7{\cdot}1$ A), and those of E seem to have *slightly* greater separation than those of F: if this apparent difference is real, it presumably corresponds to the higher ratio of soda felspar in E ($41{\cdot}2/57{\cdot}1 = 0{\cdot}72$) than in F ($35/55 = 0{\cdot}64$).

Again, we have not established with certainty the presence of exceedingly weak reflexions between the well-marked subsidiary layer-lines and the principal layer-lines, nor can we assess the significance of a tendency for subsidiary reflexions to occur in pairs, one on each of the adjacent subsidiary layer-lines.

Additional experimental data are necessary if the effects are to be fully understood, but even at this stage it is possible to advance, tentatively, suggestions as to the nature of the structures of these intermediate plagioclases.

3. STRUCTURES OF THE INTERMEDIATE PLAGIOCLASES

Just as anorthite is based on a pseudo-cell of the *same* size as the true cell of albite (linear dimensions $8 \times 13 \times 7$A, neglecting small differences), while the *true* cell is twice the size (with linear dimensions $8 \times 13 \times (2 \times 7)$ A), so the structure of any intermediate plagioclase may be supposed built up by the repetition of unit cells which are absolutely identical one with another, each cell having dimensions $8 \times 13 \times (n \times 7)$ A, where n may be large. Such a structure may be regarded as of the superlattice type, since it is based on a simple pattern with pseudo-cell 8 × 13 × 7A, but involves variations in detail which result in a larger true cell. Because of these variations, it is probable that no part of the structure is identical with either albite or anorthite.

A structure of this type offers no obvious explanation of the most characteristic feature of the X-ray photographs—the appearance of the prominent subsidiary reflexions only on the *central pairs* of subsidiary layer-lines in c-axis pictures—nor does it account naturally for other phenomena such as the subsidiary reflexions observed in a-axis and b-axis photographs. We therefore suggest the possibility of an alternative explanation, as follows.

If it is assumed that regions having the albite structure and regions having the anorthite structure coexist within the crystal, and are laid down in approximately parallel position in sheets parallel to (001), the main features of the effects observed can be explained if the dimensions of the sheets are suitably chosen. The actual calculations for the crystal have not been carried out, but our conclusions can be illustrated by means of a model in which an optical grating replaces the crystal grating. We are indebted to Professor W. L. Bragg, F.R.S., for most helpful suggestions and discussion of the details of the optical model

(i) The optical model consists of a diffraction grating in which the lines are separated by uniform intervals d, but in which the strength of the lines varies periodically so that regions in which successive lines are of nearly the same strength (F) are separated by regions in which the lines are of strengths alternately $F+2\epsilon$ and $F-2\epsilon$ to a first approximation. Thus in a region of the first kind the grating repeats at intervals d; in a region of the second kind the grating repeats at intervals $2d$ with a pseudo-repeat at intervals d if ϵ is small by comparison with F. To enable the calculations to be carried out in full, it is assumed that the strength of the nth line varies in accordance with the laws stated below, according as N is even or odd, where (Nd) is the resultant period of repeat of the grating pattern. The amplitude (A_m) of the mth order of the grating of periodicity (Nd) may be calculated from the expression

$$A_m = F_0 + F_1 \cos\frac{2\pi . m}{N} + F_2 \cos\frac{2\pi . 2m}{N} + \dots$$

$$+ F_n \cos\frac{2\pi . nm}{N} + \dots + F_{N-1} \cos\frac{2\pi . (N-1)\, m}{N},$$

where F_n represents the strength of the nth line of the grating, and is given by one or other of the following expressions:

$$N \text{ even: } F_n = F - (-1)^n \epsilon\left(1 - \cos\frac{2\pi n}{N}\right), \text{ where } n = 0, 1, 2, \ldots, (N-1),$$

$$N \text{ odd: } F_n = F + (-1)^n \epsilon\left(1 + \cos\frac{2\pi n}{N}\right), \text{ where } 0 \leqslant n < \frac{N}{2},$$

$$F_n = F - (-1)^n \epsilon\left(1 + \cos\frac{2\pi n}{N}\right), \text{ where } \frac{N}{2} < n < N.$$

In figure 2 the lengths of the lines represent the strengths F_n for simple cases of N even and odd.

The amplitudes of the spectral orders, calculated by means of the expressions given above, are as follows:

$$N \text{ even. } A_m = NF, \text{ when } m = N,$$

$$A_m = -N\epsilon, \text{ when } m = \frac{N}{2},$$

$$A_m = +\frac{N\epsilon}{2}, \text{ when } m = \frac{N}{2} \pm 1,$$

$$A_m = 0 \text{ for all other values of } m.$$

$$N \text{ odd: } A_m = NF + \epsilon \sin\frac{N\pi}{2}\left(1 - \sec\frac{\pi}{N}\right), \text{ when } m = N;$$

$$A_m = (-1)^m \epsilon \sin\frac{N\pi}{2}\left(\sec\frac{\pi m}{N} - \tfrac{1}{2}\sec\frac{\pi(m+1)}{N} - \tfrac{1}{2}\sec\frac{\pi(m-1)}{N}\right)$$

for other values of m. In all cases $A_m = A_{m+N} = A_{m+2N} =$ etc.

Thus if ϵ is small by comparison with F, when N is even we have strong principal spectra of orders N, $2N$, $3N$, ... and between each pair of consecutive principal orders a central subsidiary spectrum, bordered by two adjacent weaker spectra (one on either side), with all other spectra absent; and when N is odd, the principal spectra of orders N, $2N$, $3N$, ... have large amplitudes tending to NF as N increases, and between each pair of consecutive principal spectra there are $(N-1)$ subsidiary spectra of which the central pair, of orders $\frac{1}{2}(N-1)$ and $\frac{1}{2}(N+1)$, are much more intense than any others, as is seen by substitution for particular values of N and m. Figure 3 represents typical cases for N even and N odd. It should be noted that the Nth, $2N$th, ... orders of the grating of spacing (Nd), which we have called 'principal spectra' are the 1st, 2nd, ... orders of a perfectly uniform grating of spacing d, and the 'subsidiary spectra' of the grating (Nd) may be described as ghosts due to the periodic imperfections in the grating d.

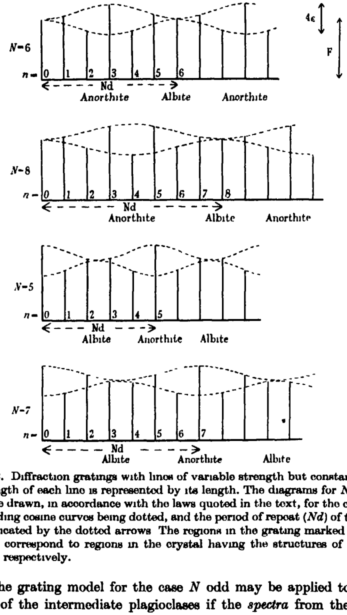


FIGURE 2. Diffraction gratings with lines of variable strength but constant spacing. The strength of each line is represented by its length. The diagrams for N even and N odd are drawn, in accordance with the laws quoted in the text, for the case $\epsilon = \frac{1}{8}F$ the bounding cosine curves being dotted, and the period of repeat (Nd) of the grating being indicated by the dotted arrows The regions in the grating marked albite and anorthite correspond to regions in the crystal having the structures of albite and anorthite respectively.

(ii) The grating model for the case N odd may be applied to the discussion of the intermediate plagioclases if the *spectra* from the grating, which has a one-dimensional periodicity, are supposed to correspond with the *layer-lines of reflexions* in *c*-axis photographs from the crystal, the comparison may be quite close, although the crystal has a three-dimensional

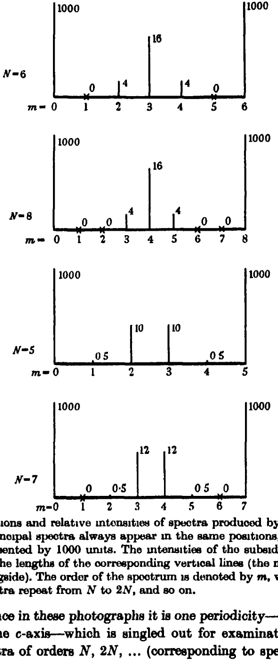


FIGURE 3. Positions and relative intensities of spectra produced by the gratings of figure 2. The principal spectra always appear in the same positions, and have large intensities represented by 1000 units. The intensities of the subsidiary spectra are represented by the lengths of the corresponding vertical lines (the numerical values are marked alongside). The order of the spectrum is denoted by m, with values from 0 to N; the spectra repeat from N to $2N$, and so on.

periodicity, since in these photographs it is one periodicity—that along the direction of the *c*-axis—which is singled out for examination. Thus the principal spectra of orders N, $2N$, ... (corresponding to spectra of orders

1, 2, ... from the ideal grating) represent the principal layer-lines (corresponding to the pseudo-axis 7·10 A), and the relatively strong central pair of 'ghost' spectra, of orders $\frac{1}{2}(N-1)$ and $\frac{1}{2}(N+1)$, represent the central pair of well-defined subsidiary layer-lines; the other fainter spectra represent the other subsidiary layer-lines the presence of which is only suspected.

Regions in the model where the grating lines at intervals d have nearly the same strength correspond to regions in the crystal which have the albite structure, with an exact repeat (along c) at intervals of 7 A; regions in the model where weaker and stronger lines alternate, so that lines of similar strengths occur at intervals $2d$, correspond to regions in the crystal with the anorthite structure which has an exact repeat (along c) at intervals of 14 A, with a well-marked pseudo-repeat of 7 A. Albite-like and anorthite-like regions are marked in figure 2.

All four materials examined—E, F, G, and Madagascar labradorite—show *pairs* of subsidiary layer-lines and so correspond to the model with N odd; we can offer no explanation at present of the apparent absence of plagioclase structures corresponding to N even, and the following discussion is therefore confined to the model for N odd.

A complete theory must account for the relationship between chemical composition and the separation of the two layer-lines of a subsidiary pair; we do not yet know this with any precision, nor do we know whether the layer-line separation varies continuously or in steps, but the model is in accordance with the experimental data so far available. Thus the layer-line separations in G, and in E and F, correspond to the positions of the spectra for models with $N=7$ and $N=5$ respectively, and there is a greater proportion of anorthite-like structure in the former, in agreement with the greater lime felspar content of specimen G. Reference to figures 2 and 3 will make this point clear. We have at present no means of knowing whether the albite and anorthite regions contain pure soda felspar and pure lime felspar respectively, or whether they contain mixed soda-lime and lime-soda felspar with compositions lying within the ranges of isomorphism near albite and anorthite respectively.

The model thus suggests a structure in which, along the direction of the c-axis, sheets of albite structure alternate with sheets of anorthite structure, each sheet being a few unit cells thick. The axes a and b of the two components will not be exactly parallel, since the interaxial angles α, β, γ are slightly different for the two structures, and the two c-axes appear to be exactly parallel. If each sheet is of considerable extent, parallel to axes a and b, the subsidiary reflexions observed in a- and b-axis photographs may correspond to the albite component, the principal reflexions to the anorthite component: in this way we can explain the abrupt though small change in dimensions of the pseudo-cell between specimen C and specimen E, and the relative weakness of subsidiary reflexions in G as compared with E and F.

4. CONCLUSION AND SUMMARY

We have determined the axial lengths of a number of plagioclases of known chemical composition ranging from nearly pure soda felspar to nearly pure lime felspar. Our observations are consistent with the existence of two separate isomorphous series (one, with the albite structure, extending from pure soda felspar inwards to at least 22% lime felspar, the other, with the anorthite structure, extending from pure lime felspar inwards to between 20 and 33% soda felspar) and of a group of intermediate plagioclases, recognizable by the characteristic arrangement of subsidiary layer-lines in *c*-axis photographs. The structural details of these three groups are discussed, and it is suggested that albite and anorthite structures coexist in the intermediate plagioclases, the features of the X-ray reflexions being reproduced in an optical analogue for which the necessary calculations can be carried to completion. Many aspects of the theory require further examination, and there is need of more experimental data.

We are indebted to Professor C. D. Jeffries, Dr E. Spencer, and Professor S. Kôzu, for the materials used in the investigation and for information, details of which are given in the text. We have also been helped by discussions, particularly on the optical analogue of the intermediate plagioclases, with Professor W. L. Bragg, F.R.S., to whom we offer our thanks. One of us (S. H. C.) is indebted to the Government of Honan Province, China, for the grant of a Scholarship during the tenure of which this work was carried out.

REFERENCES

Machatschki, F. 1928 *Zbl. Min.* **3**, 97.
Taylor, W. H., Darbyshire, J. A. and Strunz, H. 1934 *Z. Kristallogr.* **87**, 464.

The base catalysed prototropy of substituted acetones

By R. P. Bell and O. M. Lidwell

(Communicated by C. N. Hinshelwood, F.R.S.—Received 23 April 1940)

Kinetic measurements have been made at 25° on the catalysed halogenation of the following ketones: acetone, acetonylacetone, monochloroacetone, monobromoacetone, *as.* dichloroacetone and acetylacetone.

In all cases the reactions are of zero order with respect to the halogen and exhibit general basic catalysis. For the first two ketones general acid catalysis is also detectable. For each ketone the catalytic constants k_B of four carboxylate anions are related to the dissociation constants K_A of the corresponding acids by an equation of the type $k_B = G(1/K_A)^{\alpha}$, where G and α are constants for a given ketone. The catalytic constants measured cover a range of about 10^3.

Examination of the results (together with those of Pedersen for the bromination of acetoacetic ester and acetoacetic acid) shows that the exponent α decreases steadily from 0·88 to 0·48 as the reactivity of the ketone increases.

The catalytic coefficients of the water molecule are given approximately by the equations valid for the carboxylate anions, but the hydroxyl ion is in each case much less effective as a catalyst than these equations predict.

Introduction

It is now generally accepted that the catalytic effect of acids or bases in many reactions depends on a typical acid-base (protolytic) reaction involving the transfer of a proton between the catalyst and the substrate (Pedersen 1934*a*, 1938; Bell 1936, 1938). It has been found experimentally that the velocities of these reactions are correlated with the equilibrium constants of the associated acid-base equilibria: thus for catalysis by a series of similar bases* we can write

$$k_B = GK_B^{\alpha}, \tag{1}$$

where k_B is the catalytic constant, K_B a measure of the basic strength of the catalyst, and G and α are constants for a given reaction, solvent and temperature. This type of relation was first proposed by Brönsted (Bronsted and Pedersen 1924) and is usually known by his name.

Equation (1) has been interpreted in terms of the potential energy curves involved (Horiuti and Polanyi 1935; Bell 1936), the slopes of which determine the value of the exponent α. This theoretical treatment also predicts a similar dependence of the reaction velocity on the acid strength of the substrate, $(K_A)_S$. Moreover, since the potential energy curves are only linear over a restricted range, the relationship can be more properly written in the differential form

---

* The discussion given here and later refers to basic catalysis, but the same arguments apply, *mutatis mutandis*, to acid catalysis.

$$d\log k_B = \alpha\{d\log K_B + d\log (K_A)_S\}, \tag{2}$$

where α may vary continuously with k_B. By studying the variation of α experimentally it should be possible to obtain information about the potential energy curves involved. The range of velocities normally studied for a given substrate is not sufficient to reveal any appreciable variation in α: thus a thousandfold variation in k_B corresponds to a change of activation energy of only about 4 kcal./mole, while the dissociation energy of the bonds involved is of the order of 100 kcal. A very much greater variation in rate can be obtained by taking a series of related substrates, but in practice the substrate is such a weak acid that it is not possible to measure $(K_A)_S$ and thus to determine α by comparing the rates for a given catalyst and a series of substrates. However, if a number of catalysts are measured for each substrate, α is given for each substrate by $d\log k_B/d\log K_B$, and its variation over the series can be studied

In the present paper data are given for the base catalysed halogenation of compounds of the type $CH_3.CO.CHR_1R_2$. A detailed study has already been made of the effect of different catalysts on the iodination of acetone (Dawson *et al.* 1926–30) and the bromination of acetoacetic ester (Pedersen 1933, 1934*b*). Watson and Yates (1932) have made measurements on the bromination of halogen substituted acetones in 50% acetic acid which show that halogen substitution increases basic catalysis and decreases acid catalysis. They did not, however, investigate the effect of different catalytic species. It is well known that these reactions are of zero order with respect to halogen, the rate determining step involving only the ketone and the catalyst. This fact is of great practical advantage for studying a wide range of velocities. thus for a slowly reacting ketone (e g acetone) by using a very low concentration of halogen in presence of a large excess of ketone it is possible to make accurate measurements on less than 1% of the total reaction. The catalytic constants actually measured extend over a range of about nine powers of ten.

In practice complications arise in some cases owing to the substitution of successive hydrogen atoms by halogen and the reversibility of some of the halogenations. If results of the highest accuracy are needed these factors necessitate a complicated analysis of the experimental data (cf. Pedersen

1933, 1934*b*). However, the aim of the present work was to obtain a general survey of an extended series of ketones, and for this purpose a simpler treatment is adequate, the procedure adopted being described under the experimental results for each ketone. The ketones used were (in order of increasing velocity)·

Acetone	CH_3COCH_3
Acetonylacetone	$CH_3COCH_2CH_2COCH_3$
Monochloroacetone	CH_3COCH_2Cl
Monobromoacetone	CH_3COCH_2Br
as. Dichloroacetone	$CH_3COCHCl_2$
Acetylacetone	$CH_3COCH_2COCH_3$

The catalysts studied were the anions of monochloracetic, glycollic, acetic and trimethylacetic acids The basic strengths of these anions are conveniently measured by $1/K_A$, where K_A is the dissociation constant of the corresponding acid. The following values were taken for K_A at 25°:

Monochloroacetic acid	$1{\cdot}38 \times 10^{-3}$ (Wright 1934)
Glycollic acid	$1{\cdot}54 \times 10^{-4}$ (Nims 1936)
Acetic acid	$1{\cdot}75 \times 10^{-5}$ (Harned and Ehlers 1933)
Trimethylacetic acid	$9{\cdot}1 \times 10^{-6}$ (Dippy 1938)

The constants G and α of the Brönsted equation (1) were obtained in each case by applying the method of least squares to the values of $\log_{10} K_A$ and $\log_{10} k_B$ for these four anions By using suitable buffer mixtures it was also possible to deduce the catalytic effects of the bases OH^- and H_2O; in comparing these values with the Brönsted equation the dissociation constants of the corresponding acids H_2O and H_3O^+ were taken as $1{\cdot}8 \times 10^{-16}$ and $55{\cdot}5$ respectively.

EXPERIMENTAL METHOD

The halogenation velocity was measured in buffer solutions prepared by weight from sodium hydroxide solutions (free from carbonate) and acid solutions standardized directly against the sodium hydroxide. The reaction mixtures were prepared by volume from buffer solution, ketone solution, halogen solution and enough sodium chloride to make the total ionic strength of the final solution up to 0·11 (except in the monochloroacetate buffers, where the total ionic strength was 0·20). Trimethylacetic acid was prepared from pinacol hydrate (Puntambeker and Zoellner 1932) and redistilled *in vacuo*. Acetic acid was twice partly frozen, the solid being collected. Monochloroacetic acid was recrystallized twice from benzene, and trichlorophenol three times, while glycollic acid was a pure commercial specimen. All these acids were analysed by titration and found to be at least 99·7 % pure.

All kinetic measurements were carried out at 25°, and for the first four ketones the halogen used was iodine. The initial concentration of iodine was $1\text{–}2 \times 10^{-3}$ M, and the reaction was followed by pipetting out samples of 5 or 10 c.c. and titrating with N/100 thiosulphate from a micro-burette, using starch as indicator. The titres were reproducible to 0·01 c.c Eight or more titrations were carried out in each experiment. In the experiments with *as.* dichloroacetone and acetylacetone the halogen used was bromine, and the procedure adopted is described under the details for these two ketones.

Since the halogenation produces acid, the composition of the buffer solutions becomes slightly displaced during the course of the reaction. This will change the concentrations of the anion and the acid, and also the hydrogen and hydroxyl ion concentrations. It is possible to apply corrections for these changes, but in almost all cases the resulting change in the values obtained for the catalytic constants is insignificant. We have therefore in general assumed the composition of the buffer solution to remain unchanged during the reaction, though in the monochloroacetate catalysis of the acetone iodine reaction it was found necessary to apply the correction mentioned above

The rate of disappearance of halogen was found to be directly proportional to the ketone concentration when the latter was varied The velocity v given in the tables is the rate of disappearance of halogen in moles per litre per minute, referred to a ketone concentration of one mole per litre. The catalytic constants k are calculated on the assumption $v = \Sigma k_i c_i$, where c_i is the concentration of the catalyst in moles per litre.

ACETONE

The kinetics of the reaction between acetone and iodine have been extensively studied by Dawson (1926–30), but he gives values for catalysis by only two anions (acetate and monochloroacetate), and the catalytic constants given in different papers show considerable variations. Moreover, many of his experiments refer to solutions of large and varying salt concentration, and are therefore difficult to interpret quantitatively on account of primary and secondary salt effects. We have therefore reinvestigated the kinetics of this reaction.

The acetone was purified by the sodium iodide method (Shipsey and Werner 1913) and its concentration in the reaction mixtures was throughout that employed by Dawson, 0·272 M. The rate of disappearance of iodine was accurately constant throughout the course of each reaction, except in the monochloroacetate buffers, where some displacement of the buffer ratio took place during the reaction. This indicates that if more than one iodine atom is introduced, all substitutions subsequent to the first are fast compared with it. The data given later show that under conditions of basic catalysis the rates of halogenation for acetone, monochloroacetone and *as.* dichloroacetone are approximately in the ratio 1 : 400 : 3000, and the ratios for the corresponding iodoacetones are probably still greater. The rate of disappearance of iodine in presence of acetone is thus governed by the rate at which the first atom of iodine is introduced, this being followed by the rapid substitution of two further iodine atoms on the same carbon.

In a buffer solution containing c moles per litre of the basic anion B and rc moles per litre of the corresponding acid A, the total velocity is given by

$$v = v_0 + k_{OH^-}[OH^-] + k_{H_3O^+}[H_3O^+] + (k_B + rk_A)\,c,$$

where v_0 is the "spontaneous" velocity due to catalysis by water molecules. If a series of experiments is carried out with a constant buffer ratio r and a constant salt concentration, then $[OH^-]$ and $[H_3O^+]$ are constant for the series, and the above equation can be written

$$v = v' + (k_B + rk_A)\,c,$$

where $v' = v_0 + k_{OH^-}[OH^-] + k_{H_3O^+}[H_3O^+]$ is a constant for the series. The value of $(k_B + rk_A)$ can thus be obtained by plotting v against the anion concentration c, and by combining the results of several such series with different values of r the values of k_B and k_A are obtained. This method of analysis has been applied to the results for acetone and acetonylacetone. The value of v' (obtained by extrapolating v to zero buffer concentration) is given for each series, and the calculated values of v are obtained from v' and the values of k_B and k_A given at the head of each section. The concordance of the observed and calculated values of v shows the applicability of the above equations.

In the case of the monochloroacetate buffers there is a slight increase in the rate of iodine consumption during the course of each reaction, indicating an appreciable change in the buffer composition. The velocities given in the table are mean values. The mean buffer ratio is easily calculated

TABLE 1. IODINATION OF ACETONE

Trimethylacetate. $10^6 k_B = 24·5$, $10^6 k_A = 4·4$

$r = 0·490$, $10^7 v' = 0·74$.

$10^2 c$	20·3	15·2	10·1	5·07
$10^7 v$ Obs.	6·07	4·78	3·35	2·17
Calc.	6·13	4·78	3·41	2·10

$r = 1·015$, $10^7 v' = 0·81$.

$10^2 c$	20·2	15·1	10·1	5·05
$10^7 v$ Obs.	6·61	5·15	3·64	2·35
Calc.	6·65	5·18	3·75	2·28

Acetate. $10^6 k_B = 14·6$, $10^6 k_A = 6·06$

$r = 0·210$, $10^7 v' = 0·63$.

$100 c$	10·0	7·50	5·00	2·50
$10^7 v$ Obs.	16·6	12·5	8·57	4·64
Calc.	16·8	12·6	8·58	4·61

$r = 0·470$, $10^7 v' = 0·52$.

$100 c$	10·0	7·50	5·00	2·50
$10^7 v$ Obs.	18·1	13·7	9·15	4·85
Calc.	18·0	13·6	9·25	4·85

$r = 1·11$, $10^7 v' = 0·81$.

$100 c$	10·0	7·50	5·00	2·50
$10^7 v$ Obs.	22·2	17·9	11·4	6·19
Calc.	22·0	17·7	11·4	6·11

Glycollate. $10^6 k_B = 2·31$, $10^6 k_A = 19·5$

$r = 0·0840$, $10^7 v' = 0·85$.

$100 c$	10·0	7·50	5·00	2·50
$10^7 v$ Obs.	4·78	3·90	2·76	1·91
Calc.	4·81	3·81	2·83	1·84

$r = 0·201$, $10^7 v' = 1·22$.

$100 c$	10·0	7·50	5·00	2·50
$10^7 v$ Obs.	7·50	5·91	4·30	2·83
Calc.	7·48	5·91	4·34	2·80

Monochloroacetate. $10^6 k_B = 0·30$, $10^6 k_A = 76$

Mean ratio = 0·0260, $10^7 v' = 0·92$.

$100 c$	19·0	14·2	9·48	4·74
$10^7 v$ Obs.	5·15	4·30	3·25	2·36
Corr.	5·37	4·35	3·16	1·95
Calc.	5·26	4·16	3·08	2·00

Mean ratio = 0·0376, $10^7 v' = 1·42$.

$100 c$	19·0	14·3	9·50	4·75
$10^7 v$ Obs.	7·06	5·60	4·38	3·01
Corr.	7·40	5·85	4·38	2·87
Calc.	7·40	5·92	4·41	2·90

Mean ratio = 0·0535, $10^7 v' = 2·39$.

$100 c$	19·0	14·2	9·48	4·74
$10^7 v$ Obs.	10·2	8·20	6·48	4·60
Corr.	10·6	8·25	6·48	4·42
Calc.	10·7	8·59	6·53	4·46

for each experiment, but in a series with equal initial ratios the mean ratios will not in general be equal for different buffer concentrations. Since the catalysis by monochloroacetate ions is small compared with that by monochloroacetic acid, it is necessary to correct for this inequality in order to obtain a reliable value for the former. The values of v (corr.) in the table represent the velocity which would have been observed if the actual mean buffer ratio in each experiment of a series had been equal to the average value over all the experiments of that series. The correction is readily calculated from approximate values of the catalytic constants of OH^-, H_3O^+ and monochloroacetic acid. The corrected values can then be analysed as described above.

The buffer solutions in table 1 are not very suitable for deriving the catalytic constants of hydroxyl ion and water, and additional measurements were made in a trichlorophenate buffer of ratio approximately 0·08. The dissociation constant of trichlorophenol is not known with any accuracy, and the pH of the buffer solution was found to be 6·84 by direct measurement with a glass electrode. Trichlorophenate buffers react slowly with iodine even after thorough purification. In order to correct for this the initial rates of iodination were measured in the absence of acetone and were subtracted from the observed rates. Catalysis by undissociated trichlorophenol will certainly be negligible in these solutions

TABLE 2

Trichlorophenate. $10^6 k_B = 454$

$r =$ *ca.* $0{\cdot}08$, pH $= 6{\cdot}84$, $10^7 v' = 13{\cdot}6$.

$10^3 c$	22·0	16·5	11·0	5·5
$10^6 v$ Obs.	14·9	11·6	8·10	4·98
Corr	11·5	8·8	6·23	3·93
Calc.	11·4	8·9	6·35	3·85

The catalytic constant of the *hydrogen ion* is given by the measurements of Dawson and Powis (1913) on solutions of strong acids. In the present units the value is $10^6 k_A = 1640$.

It is now possible to use the values of v' in the buffer mixtures to obtain the catalytic constants of the *hydroxyl ion* and the *water molecule*. In addition to the above values we have recalculated the data of Dawson and Key (1928) for acetate and phosphate buffers. In calculating the hydrogen and hydroxyl ion concentrations from the thermodynamic dissociation constants we have assumed that the activity coefficients of all univalent ions are given by

$$-\log_{10} f_{\pm} = \frac{0{\cdot}50\sqrt{\mu}}{1+\sqrt{\mu}},$$

where μ is the ionic strength (cf. Guggenheim 1935), and that the activity coefficients of all neutral molecules are unity.† In the case of the trichlorophenate buffer, $[H_3O^+]$ and $[OH^-]$ are calculated from the measured pH, making the same assumptions about the activity coefficients. The calculated values in table 3 are given by

$$10^7\, v' = 5 \times 10^{-3}\,[H_2O] + 1{\cdot}64 \times 10^4\,[H_3O^+] + 1{\cdot}5 \times 10^6\,[OH^-],$$

where $[H_2O] = 55{\cdot}5$ throughout.

TABLE 3. CATALYSIS BY H_3O^+, OH^- AND H_2O

Buffer	Ratio	$10^9[OH^-]$	$10^7 v'$ (obs.)	$10^7 v'$ (calc.)
*Acetate	3·00	0·19	1·3	1·3
,,	1·11	0·51	0·8	0·8
* ,,	1·00	0·57	0·8	0·7
,,	0·47	1·21	0·5	0·6
* ,,	0·33	1·71	0·8	0·7
,,	0·21	2·70	0·6	0·8
* ,,	0·10	5·7	1·6	1·2
Trimethylacetate	1·02	1·07	0·8	0·6
,,	0·49	2·20	0·7	0·8
*Phosphate	0·21	33·7	5·0	5·3
Trichlorophenate	0·08	90	13·6	13·6

* Results from Dawson and Koy (1928).

The agreement between observed and calculated values is as good as could be expected, since most of the values of v' represent a small fraction of the observed velocities. The value of k_{OH^-} is only about half as great as that obtained by Dawson and Koy (1928), but there appears to be a miscalculation in their value for the pH of the phosphate buffer, upon which their value of k_{OH^-} largely depends. The values of k_{OH^-} and k_{H_2O} are of course known much less accurately than the other catalytic constants.

The catalytic constants for acetone are summarized in table 4, which also gives the values calculated from the Brönsted equation and those reported by Dawson. The latter are given differently in different papers, the limits of variation being given in the table.

ACETONYLACETONE

A commercial sample of the ketone was twice fractionated under reduced pressure, the final product boiling within 0·2° at 99°/30 mm. The concentration of the ketone in the reaction mixtures was 0·189 M throughout.

---

† In the buffers used by Dawson the ionic strength was not constant, and we have taken the mean value for each buffer ratio. The second dissociation constant of phosphoric acid has been taken as $6{\cdot}2 \times 10^{-8}$ (Bjerrum and Unmack 1929).

With the exception of the monochloroacetate buffers (where the buffer ratio is slightly displaced) the rate of disappearance of iodine was constant throughout each experiment. As in the case of acetone, this shows that the substitution of the first atom of iodine is slow compared with any subsequent substitutions. The chief reaction taking place will therefore be the formation of $CH_3COCH_2CI_2COCH_3$. However, under conditions of basic catalysis the rate of iodination is only about six times as fast as for acetone, and it is therefore likely that iodination of the terminal methyl groups also contributes appreciably to the observed velocity. From a theoretical point of view it would be desirable to separate the effects due to the two groups, and this might be done approximately by assuming that the terminal groups react at the same rate as those in acetone. Actually, however, this correction has little effect on the exponent of the Brönsted equation, and for the purposes of the present paper we shall use the total reaction velocities without attempting to separate the contributions of the different groups.

TABLE 4. CATALYTIC CONSTANTS FOR THE IODINATION OF ACETONE

Basic catalysis. Calculated values from $k_B = 9{\cdot}1 \times 10^{-10}(1/K_A)^{0\cdot88}$

	$10^6 k_B$		
Catalyst	Obs	Calc	Dawson
CH_2ClCOO	0·30	0·30	0 44
CH_2OHCOO^-	2·3	2·1	—
CH_3COO^-	14·6	14·7	12·1–16·5
CMe_3COO^-	24·5	24·4	—
$C_6H_3Cl_2O^-$	450	400	—
OH^-	$1{\cdot}5 \times 10^7$	$5\ 2 \times 10^{10}$	$2{\cdot}6 \times 10^7$
H_2O	5×10^{-4}	$1{\cdot}8 \times 10^{-5}$	5×10^{-4}

Acid catalysis. Calculated values from $k_A = 2{\cdot}4 \times 10^{-3} K_A^{0\cdot88}$

	$10^6 k_A$		
Catalyst	Obs.	Calc.	Dawson
$CH_2ClCOOH$	76	68	82–90
$CH_2OHCOOH$	19·4	20·5	20 8
CH_3COOH	6·1	6·2	4·9-5·9
CMe_3COOH	4·4	4 4	—
H_2O	—	$2{\cdot}2 \times 10^4$	$1{\cdot}64 \times 10^3$

The treatment of the results is in most cases the same as described for acetone. In the case of the monochloroacetate buffers, however, a simpler procedure is possible, since the ratio k_A/k_B is now about 40 instead of 250. The rates given correspond to the first half of the reaction, and although the mean buffer ratio in each series is appreciably different from the initial ratio, it is not necessary to correct for the differences in buffer ratio between

TABLE 5. IODINATION OF ACETONYLACETONE

Trimethylacetate. $10^6 k_B = 151$

$r = 0\cdot220$, $10^7 v' = 6\cdot5$.

$100c$	5·30	3·96	2·66	1·32
$10^7 v$ Obs.	87·5	67·7	45·9	23·9
Calc.	87·1	66·6	46·9	25·6

Acetate. $10^6 k_B = 85\cdot6$, $10^6 k_A = 3\cdot8$

$r = 0\cdot289$, $10^7 v' = 3\cdot7$.

$100c$	10·27	7·71	5·14	2·57
$10^7 v$ Obs.	92·8	71·2	47·6	26·0
Calc.	92·7	70·6	47·2	26·0

$r = 0\cdot977$, $10^7 v' = 2\cdot4$.

$100c$	9·92	7·45	4·96	2·48
$10^7 v$ Obs.	90·3	69·1	47·2	24·3
Calc.	90·8	68·9	46·7	24·6

Glycollate. $10^6 k_B = 17\cdot6$, $10^6 k_A = 13$

$r = 0\cdot106$, $10^7 v' = 2\cdot5$.

$100c$	10·3	7·74	5·15	2·58
$10^7 v$ Obs.	22·1	17·0	12·2	7·5
Calc.	22·2	17·1	12·4	7·4

$r = 0\cdot138$, $10^7 v' = 2\cdot8$.

$100c$	10·3	7·73	4·72	2·58
$10^7 v$ Obs.	23·0	17·4	11·9	8·5
Calc.	23·0	18·0	12·1	7·9

Monochloroacetate. $10^6 k_B = 1\cdot5$, $10^6 k_A = 58$

Mean ratio $= 0\cdot0424$, $10^7 v' = 3\cdot1$.

$100c$	17·8	13·3	8·88
$10^7 v$ Obs	9·9	8·2	6·8
Calc.	10·1	8·3	6·6

Mean ratio $= 0\cdot0902$, $10^7 v' = 4\cdot5$.

$100c$	18·4	13·8	9·20
$10^7 v$ Obs.	17·1	14·0	10·7
Calc.	16·9	13·8	10·7

Mean ratio $= 0\cdot142$, $10^7 v' = 6\cdot3$.

$100c$	18·5	13·9	9·25	4·63
$10^7 v$ Obs.	25·0	21·1	16·1	12·0
Calc.	26·2	21·3	16·2	11·3

Hydrogen ion (HCl solutions). $10^6 k_A = 1510$

$100c$	1·00	1·92	3·72
$10^7 v$ Obs.	145	308	559
Calc.	151	290	562

the single experiments of a series. Further, it is clear from the results for other acids that the catalytic power of trimethylacetic acid will be very small compared with that of the trimethylacetate ion. Only one buffer ratio was therefore investigated for this catalyst, acid catalysis being neglected.

The data for acetonylacetone are not well adapted for calculating k_{OH^-} or k_{H_2O}, but rough values can be obtained from the values of v' in table 5. These are collected in table 6, the calculated values being given by

$$10^7 v' = 2{\cdot}5 \times 10^{-3}[H_2O] + 1{\cdot}0 \times 10^9[OH^-] + 1{\cdot}51 \times 10^4[H_3O^+].$$

The hydrogen and hydroxyl concentrations were calculated as described under acetone.

TABLE 6. CATALYSIS BY H_3O^+, OH^- and H_2O

Buffer	Ratio	$10^5[H_3O^+]$	$10^{10}[OH^-]$	$10^7 v$ Obs.	$10^7 v$ Calc.
Trimethylacetate	0·22	0·31	51	6·5	6·5
Acetate	0·29	0·79	20	3·7	3·6
,,	0·98	2·7	5·9	2·4	2·4
Glycollate	0·11	2·5	6·4	2·5	2·4
,,	0·14	3·4	4·7	2·8	2·4
Monochloroacetate	0·042	9·9	1·6	3·1	3·1
,,	0·090	21	0·8	4·5	4·3
,,	0·14	33	0·5	6·3	6·2

The collected results for acetonylacetone are given in table 7. The acid catalytic constants are not very accurate, but it will be seen that they differ little from those for acetone.

TABLE 7 CATALYTIC CONSTANTS FOR THE IODINATION OF ACETONYLACETONE

Calculated values from $k_B = 5{\cdot}24 \times 10^{-9}(1/K_A)^{0{\cdot}89}$

Catalyst	$10^6 k_B$ (obs.)	$10^6 k_B$ (calc.)	$10^5 k_A$
CH_2ClCOO^-	1·5	1·8	58
CH_2OHCOO^-	17·6	12·9	13
CH_3COO^-	86	89	3·8
CMe_3COO^-	151	160	—
OH^-	$1{\cdot}0 \times 10^8$	$5{\cdot}5 \times 10^{11}$	—
H_2O	$2{\cdot}5 \times 10^{-3}$	$1{\cdot}5 \times 10^{-4}$	$1{\cdot}51 \times 10^3$

MONOBROMOACETONE

The ketone was prepared by the direct bromination of acetone (Levene 1930) and was twice fractionated under reduced pressure. The final product boiled at 50–51°/15 mm. and was colourless when freshly distilled. On standing it gradually became brown and the rate of iodination increased. All the results given here were obtained with freshly distilled specimens.

The reaction mixtures used had ketone concentrations (m) between 0·02 and 0·06 M, and concordant results were obtained with different concentrations. The rate of disappearance of iodine was not constant throughout each experiment, a typical curve being given in figure 1 A. The first part of this curve shows an acceleration up to a steady rate, which is maintained for the greater part of the reaction. This suggests that the iodination involves two consecutive reactions each of the first order with respect to ketone and of zero order with respect to iodine. If $[S]$ is the concentration of ketone (effectively constant during the reaction) and k_1 and k_2 the first order

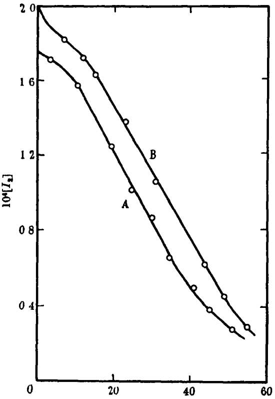


FIGURE 1

constants for the two reactions, then it is easily shown that the rate of disappearance of iodine is given by

$$-\frac{d[I_2]}{dt} = k_1[S](2-e^{-k_2t}).$$

If k_2 is considerably greater than k_1 this expression reduces to

$$-d[I_2]/dt = 2k_1[S]$$

for moderately large values of t, corresponding to the linear part of the observed curve. (If k_1/k_2 is very great, the initial acceleration will be undetectable, as in the case of the first two ketones.) The falling off of the velocity towards the end of the reaction suggests that the second iodination does not go quite to completion, and in fact a small amount of iodine always remained in equilibrium at the end of the reaction. This reversibility might interfere with the titration of iodine, but experiments in which the disappearance of iodine was followed colorimetrically gave essentially the same curves and were much more laborious to carry out. The titration method was therefore used, and the velocities recorded correspond to the linear parts of the reaction curves. Any slight errors involved in this procedure will be the same for different catalysts, and are hence unimportant for the purposes of the present paper. The velocities involved are several hundred times greater than those for acetone under the same conditions, so that the reaction of the unsubstituted methyl group can be neglected. The same applies to the still more reactive ketones dealt with in the following sections.

In solutions of strong acids the iodination becomes reversible to such an extent that only very rough estimates of the reaction rate can be made. It is, however, clear that the catalytic constant of the hydrogen ion is less than 2×10^{-4} (i.e. less than that of the monochloroacetate ion), and it can be concluded that catalysis by undissociated carboxylic acid molecules will be much too small to contribute to the rate in buffer solutions. This simplifies the procedure for obtaining the catalytic constants of the anions, since it is necessary to use only one buffer ratio for each catalyst (cf. also the results for monochloroacetone).

Approximate values of k_{OH^-} and k_{H_2O} can be obtained as before from the values of v'. The calculated values in table 9 are given by

$$10^5 v = 2\cdot9\times10^{-2}[H_2O]+1\cdot2\times10^{9}[OH^-].$$

The collected catalytic constants for monobromoacetone are given in table 10.

TABLE 8. IODINATION OF MONOBROMOACETONE

Trimethylacetate. $10^4 k_B = 158$, $r = 0\cdot078$, $10^5 v' = 19$

$100c$	5·00	3·75	2·50
$100m$	1·92	1·92	3·85
$10^5 v$ Obs.	98	78	58
Calc.	98	78	58

Acetate. $10^4 k_B = 112$, $r = 0\cdot0805$, $10^5 v' = 10$

$100c$	10·00	7·50	5·00	2·50
$100m$	3·85	3·85	3·85	3·85
$10^5 v$ Obs.	121	95	68	35
Calc.	122	94	66	38

Glycollate. $10^4 k_B = 15\cdot1$, $r = 0\cdot083$, $10^5 v' = 2\cdot2$

$100c$	10·00	7·50	5·00	2·50
$100m$	3·85	3·85	5·77	5·77
$10^6 v$ Obs.	176, 175	126, 133	94, 95	60, 58, 59
Calc.	173	135	97	60

Monochloroacetate. $10^4 k_B = 2\cdot75$, $r = 0\cdot0235$, $10^5 v' = 2\cdot4$

$100c$	19·0	14·3	9·50	4·75
$100m$	3·85	3·85	5·77	5·77
$10^5 v$ Obs.	7·8	6·0	5·4	3·8
Calc.	7·6	6·3	5·0	3·7

TABLE 9. CATALYSIS BY OH^- AND H_2O

Buffer	Ratio	$10^9[OH^-]$	$10^5 v'$ (obs.)	$10^5 v'$ (calc.)
Trimethylacetate	0·078	14·4	19	19
Acetate	0·081	7·1	10	10
Glycollate	0·083	0·79	2·2	2·5
Monochloroacetate	0·024	0·32	2·4	2·0

TABLE 10. CATALYTIC CONSTANTS FOR THE IODINATION OF MONOBROMOACETONE

Calculated values from $k_B = 1\cdot18 \times 10^{-6} (1/K_A)^{0\cdot82}$

Catalyst	$10^4 k_B$ (obs.)	$10^4 k_B$ (calc.)
CH_2ClCOO^-	2·75	2·68
CH_2OHCOO^-	15·1	16·3
CH_3COO^-	112	98
CMe_3COO^-	158	168
OH^-	$1\cdot2 \times 10^8$	$1\cdot0 \times 10^{11}$
H_2O	$2\cdot9 \times 10^{-8}$	$4\cdot4 \times 10^{-4}$

MONOCHLOROACETONE

Difficulties were experienced in purifying this ketone. A commercial sample was fractionated three times through a 45 cm. column, giving a pale yellow product boiling at 118·5–118·9°. It still contained about 2 % of an unsaturated product (possibly mesityl oxide) which reacted immediately with bromine, but not with iodine. Further fractionation caused no noticeable improvement.

The concentration of ketone in the reaction mixtures varied from 0·005 M to 0·2 M. The course of the iodination (curve *B*, figure 1) was very similar to that for monobromoacetone, though the linear portion extended over a larger fraction of the reaction. There was, however, a rapid disappearance of a small quantity of iodine during the first few minutes. This could be accounted for by the presence of about 0·1 % of a more reactive ketone, and it may be noted that there is only 1° difference between the boiling points of monochloroacetone and *as*. dichloroacetone, a likely impurity. The initial fall is not likely to be due to the presence of a small quantity of the enol form of monochloroacetone, since no initial drop was found for monobromoacetone or *as*. dichloroacetone.

The velocities given in table 11 are taken from the linear part of the reaction curves (amounting to about 60 % of the reaction) and will be little affected by any of the complicating factors mentioned. This applies particularly to the relative velocities for different catalysts. The results in solutions of hydrochloric acid show that the hydrogen ion is not a very effective catalyst, being considerably less active than the acetate ion. It is therefore likely that catalysis by undissociated carboxylic acid molecules can be neglected, and this is confirmed by the results in the acetate and monochloroacetate buffers of different ratios. In the trichlorophenate buffers the results have been corrected (cf. under acetone) for the rate of iodination of the buffer solution itself.

The values of v' in table 11 are well represented by

$$10^5 v' = 5{\cdot}8\times10^{-3}[\mathrm{H_2O}]+5{\cdot}6\times10^{8}[\mathrm{OH^-}]$$

the effect of the hydrogen ion being negligible. Table 12 shows the comparison between observed and calculated values.

The collected catalytic constants are given in table 13. It will be seen that the value for the trichlorophenate ion does not agree with that calculated from the results for the carboxylate anions. There is considerable uncertainty about the dissociation constant of trichlorophenol (Tiessens 1929, Ogston 1936): however, the value of $3{\cdot}9\times10^{-7}$ used here gave good

TABLE 11. IODINATION OF MONOCHLOROACETONE

Trimethylacetate. $10^4 k_B = 111$, $r = 0{\cdot}078$, $10^5 v' = 8{\cdot}7$

$100c$	4·97	3·72	2·44	1·19
$100m$	5·45	5·45	10·9	10·9
$10^5 v$ Obs.	64	50	36	22
Calc.	64	50	36	22

Acetate. $10^4 k_B = 64{\cdot}3$

$r = 0{\cdot}281$, $10^5 v' = 1{\cdot}5$.

$100c$	10·00	7·50	5·00	2·50
$100m$	5·45	5·45	10·9	10·9
$10^5 v$ Obs.	68	50	36	17
Calc.	66	50	34	18

$r = 2{\cdot}11$, $10^5 v' = 0{\cdot}4$.

$100c$	10·00	7·50	5·00	2·50
$100m$	5·45	5·45	10·9	10·9
$10^5 v$ Obs.	63	48	33	16
Calc.	65	49	33	16

Glycollate. $10^4 k_B = 10{\cdot}5$, $r = 0{\cdot}084$, $10^5 v' = 0{\cdot}8$

$100c$	10·00	7·50	5·00
$100m$	10·9	10·9	21·8
$10^5 v$ Obs.	11·2	8·2	6·4
Calc.	11·3	8·7	6·1

Monochloroacetate. $10^4 k_B = 1{\cdot}78$

$10^5 v' = 0{\cdot}5$, (*a*) $r = 0{\cdot}0235$, (*b*) $r = 0{\cdot}0470$.

$100c$	19·0	14·3	9·50	4·75	4·75	2·38
$100m$	10·9	10·9	10·9	10·9	16·4	16·4
$10^5 v$ Obs. (*a*)	3·9	2·9	2·1	1·3	1·4	1·0
Obs. (*b*)	3·9	3·0	2·2	1·3	1·4	1·0
Calc.	3·9	3·0	2·2	1·35	1·35	0·9

Trichlorophenate. $10^4 k_B = 3180$, $r = ca$ $0{\cdot}08$, $10^5 v' = 52$
pH = 6·84 (glass electrode)

$100c$	2·20	1·76	1·10	0·44
$100m$	0·545	0·545	1·09	1·09
$10^4 v$ Obs.	77	64	40	20
Corr.	74	62	39	20
Calc.	76	61	40	19

Hydrogen ion (HCl solutions). $10^4 k_A = 37{\cdot}9$, $10^5 v' = 0{\cdot}3$

$100c$	4·40	3·53	1·81
$100m$	10·9	10·9	21·8
$10^5 v$ Obs.	17·3	13·4	7·2
Calc.	17·0	13·7	72·2

agreement in the iodination of acetone (cf. table 4), so that there is a real discrepancy.

TABLE 12. CATALYSIS BY OH^- AND H_2O

Buffer	Ratio	$10^9[OH^-]$	$10^5 v'$ (obs.)	$10^5 v'$ (calc.)
Trichlorophenate	0 08	90	52	51
Trimethylacetate	0 078	14·4	8·7	8 9
Acetate	0·28	2 0	1·5	1·3
,,	2 11	0 27	0 4	0·4
Glycollate	0·084	0·77	0 8	0 8
Monochloroacetate	0·0235	0 32	0 5	0·5
,,	0·0470	0·16	0 5	0 4
HCl	—	0·0	0 3	0·3

TABLE 13. CATALYTIC CONSTANTS FOR THE IODINATION OF MONOCHLOROACETONE

Calculated values from $k_B = 8\ 04 \times 10^{-7}(1/K_A)^{0·82}$

Catalyst	$10^4 k_B$ (obs.)	$10^4 k_B$ (calc.)
CH_2ClCOO^-	1·78	1·78
CH_2OHCOO^-	10 5	10 8
CH_3COO	64	64
CMe_3COO	111	109
$C_6H_2Cl_3O$	$3·2 \times 10^3$	$1·5 \times 10^3$
OH^-	$5·6 \times 10^6$	$1\ 5 \times 10^{10}$
H_2O	$5·8 \times 10^{-8}$	$3\ 0 \times 10^{-8}$
H_3O^+	$10^3 k_A = 3\ 79$	—

as. DICHLOROACETONE

The ketone (b p. 115–116°) was prepared by hydrolysing α-dichloroacetoacetic ester with 5 N hydrochloric acid Its chlorine content was determined by warming a sample with 2 N sodium hydroxide, neutralizing, and titrating the chloride ion with silver nitrate solution, the end-point being determined electrometrically. (Found 55·7 % Cl, calculated 55·9 %.)

The iodination of *as.* dichloroacetone is a reversible reaction, a large amount of iodine being present at equilibrium. It was therefore necessary to use the rate of bromination as a measure of the rate of ionization of this ketone. Preliminary experiments showed that the reaction is of zero order with respect to bromine, and that no free bromine remains at equilibrium. The reaction presents, however, a number of complications, which have been studied in detail by Mr D. H. Everett in this laboratory. Only the main features will be mentioned here, and the catalytic constants obtained are less certain than for the other ketones. The titration method is subject to some error owing to the loss of bromine by evaporation, especially in experiments of long duration. A number of experiments were therefore carried out in which a known amount of bromine was added and the time taken for its disappearance observed by measuring the potential of a platinum electrode in the solution. These experiments are marked with an asterisk in table 14. Abnormally high rates were obtained in the more alkaline buffers, possibly owing to oxidation by hypobromite, and the

TABLE 14. BROMINATION OF *as.* DICHLOROACETONE

Trimethylacetate. $10^2 k_B = 8{\cdot}06$

$r = 0{\cdot}130$, $10^4 v' = 2{\cdot}7$.

100 c	1·50	2·98	3·00	3·00	3·00
100 m	1·05	1·08	1·05	1·05	2·17
$10^4 v$ Obs.	15·7	23·2	26·7	25·6	26·8
Calc.	14·8	26·7	26·8	26·8	26·8

100 c	4·47	4·50	5·96	6·00
100 m	1·04	1·05	1·04	1·05
$10^4 v$ Obs.	34·6	39·3	47·4	50·8
Calc.	38·7	39·0	50·7	51·1

Acetate. $10^2 k_A = 3{\cdot}38$

$r = 0{\cdot}487$, $10^4 v' = 0{\cdot}8$.

100 c	2·50	5·00	7·50
100 m	0·75	0·50	0·30
$10^4 v$ Obs.	9·8*	17·5*	26·5*
Calc.	9·3	17·7	26·2

$r = 2{\cdot}00$, $10^4 v' = 0{\cdot}5$

100 c	2·00	2·50	2·50	4·00	5·00	5·00
100 m	1·05	1·08	0·30	1·05	1·08	0·30
$10^4 v$ Obs.	7·3	9·2	9·8*	13·2	17·9	17·5*
Calc.	7·3	9·0	9·0	13·9	17·4	17·4

100 c	6·00	7·50	7·50	8·00	8·00	8·00
100 m	1·05	1·08	0·30	0·50	1·05	2·19
$10^4 v$ Obs	20·4	26·6	26·5*	29·3	26·4	25·6
Calc.	20·8	25·9	25·9	27·6	27·6	27·6

Monochloroacetate $10^2 k_B = 0{\cdot}125$

$r = 1{\cdot}11$, $10^4 v' = 0{\cdot}44$.

100 c	2·50	5·00	7·50	7·50
100 m	1·60	1·60	1·60	1·60
$10^4 v$ Obs.	0·76	1·09	1·33	1·37
Calc.	0·75	1·07	1·38	1·38

Hydrochloric acid. $10^4 k_A (?) = 0{\cdot}11$, $10^4 v' = 0{\cdot}437$

100 c	9·4	37·6	119
100 m	2·59	2·05	2·04
$10^4 v$ Obs.	0·449	0·475	0·570
Calc.	0·447	0·478	0·568

recorded velocities refer in most cases to buffers of fairly high ratio. This was not feasible for trimethylacetate, but in this case the velocity of the catalysed reaction is so great that any correction will be small. It was also found that the rate increased when successive quantities of bromine were added to the same reaction mixture, possibly owing to rearrangement of the brominated ketone, and the recorded values refer to the early stages of the reaction. It was not possible to use glycollate buffers, since these react with bromine at a rate which is comparable to that of the ketone bromination.

The increase of rate observed in acid solution is so small that it may well be due to a salt effect rather than to acid catalysis. In any case it is clear that catalysis by undissociated carboxylic acid molecules can be neglected, and this is confirmed by the concordant results for acetate buffers of different ratios The catalytic constant of the water molecule is given by the results in acid solution as $4\cdot4 \times 10^{-5}/55\cdot5 = 7\cdot9 \times 10^{-7}$. The only information as to catalysis by hydroxyl ion comes from the value of v' in the trimethylacetate buffer, which gives $k_{OH^-} = 2\cdot7 \times 10^4$. The catalytic constants for *as.* dichloroacetone are collected in table 15

TABLE 15. CATALYTIC CONSTANTS FOR THE BROMINATION OF *as* DICHLOROACETONE

Calculated values from $k_B = 7\cdot2 \times 10^{-6}(1/K_A)^{0\cdot78}$

Catalyst	$10^2 k_B$ (obs.)	$10^2 k_B$ (calc.)
CH_2ClCOO^-	0 125	0·125
CH_3COO^-	3 4	3·7
CMe_3COO^-	8·1	6 3
OH^-	$2\cdot7 \times 10^6$	$1\ 4 \times 10^9$
H_2O	$7\cdot9 \times 10^{-5}$	$3\cdot0 \times 10^{-5}$
H_3O^+	$10^4 k_A(?) = 0\ 11$	-

ACETYLACETONE

A commercial sample was redistilled and boiled at 136°. The titration methods could not be used, since the reaction is very fast even with dilute ketone solutions, and the halogenation is slightly reversible even when bromine is used. However, the bromoketone formed will liberate iodine from potassium iodide, and the following procedure was found to give satisfactory results.

9·1 c c. of 4×10^{-4} M ketone solution was brought to 25° and 0·9 c.c. of bromine water (also at 25°) was added rapidly with shaking, using the type of pipette previously described (Bell, Lidwell and Vaughan-Jackson 1936). After a suitable interval the reaction was stopped by adding rapidly 2 c.c of M/10 allyl alcohol, which removes the remaining free bromine. The amount of bromo-ketone formed was then determined by adding acidified potassium iodide solution and titrating the liberated iodine. With practice it was possible to obtain consistent results with a reaction time as short as 3 sec.

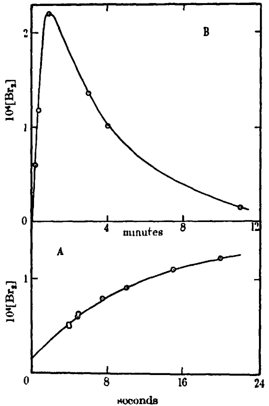


FIGURE 2

Typical reaction curves are shown in figure 2. Curve *A* represents the beginning of the reaction and curve *B* its behaviour over longer periods. It will be seen that the bromo-ketone slowly changes into some other product which will not liberate iodine from potassium iodide. (The same behaviour is shown by α-dibromoacetoacetic ester cf. Pedersen 1934*b*.) This change will be negligible in the early part of the reaction, which was used for determining the rate of bromination. Since the ketone concentration is small it is necessary to allow for its variation during the course of the reaction. If m is the initial concentration of ketone and x the iodine concentration (in moles per litre) at any time, then in the absence of any disturbing factors

$\log_{10}(m-\frac{1}{2}x)$ should vary linearly with the time. Figure 3 shows that this is so for the early stages of the reaction. The subsequent falling off is partly due to the reversible nature of the bromination, and partly to the slow transformation of the dibromoketone. The rate of bromination in moles of bromine per mole of ketone per litre is then given by

$$\frac{2\times 2{\cdot}303}{m}\,\frac{d\log_{10}(m-\frac{1}{2}x)}{dt}.$$

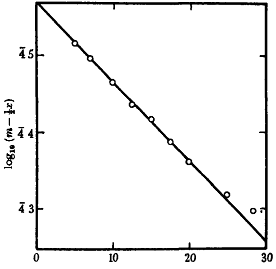


FIGURE 3

At the beginning of each reaction a certain amount of bromine appeared to be used up instantaneously (see the intercepts at $t=0$ in figures 2A and 3). This corresponded throughout to 18 % ± 2 % of the ketone present, and is no doubt due to the fraction which is present as enol. Earlier measurements (Meyer 1911) give 19 % for the equilibrium proportion of enol in aqueous solutions at 0°.

The results in solutions of hydrochloric acid show that catalysis by hydrogen ions is negligible, and give an accurate value for the "spontaneous" reaction. Moreover, the values of v' obtained in the buffer solutions are identical with the spontaneous rate within the experimental error, showing that hydroxyl ion catalysis is also negligible in the solutions used. From this it can be concluded that $k_{OH^-} < 10^7$.

The catalytic constants for acetylacetone are collected in table 17.

TABLE 16. THE BROMINATION OF ACETYLACETONE

Ketone concentration $= 4{\cdot}0 \times 10^{-4}$ throughout. $v' = 1{\cdot}36 = 0{\cdot}024 \times 55{\cdot}5$ throughout

	Trimethylacetate. $k_B = 508$, $r = 0{\cdot}0780$			
$100c$	0·875	0·625	0·375	
v (obs.)	5·99	4·42	3·25	
v (calc.)	5·82	4·54	3·25	
	Acetate. $k_B = 308$, $r = 0{\cdot}0805$			
$100c$	2·50	1·25	0·500	
v (obs.)	8·67	5·01	2·74	
v (calc.)	8·88	5·01	2·70	
	Glycollate $k_B = 94$, $r = 0{\cdot}0830$			
$100c$	5·00	3·75	2·50	1·25
v (obs.)	6·28	4·97	3·59	2·50
v (calc.)	6·06	4·88	3·72	2·54
	Monochloroacetate $k_B = 36$, $r = 0{\cdot}0470$			
$100c$	9·48	7·11	4·74	2·37
v (obs.)	4·84	3·79	3·04	2·19
v (calc.)	4·76	3·92	3·06	2·22
	Hydrochloric acid. $k_A = 0$			
$100c$	37·6	9·40	0·94	
v (obs.)	1·38	1·36	1·36	

TABLE 17. CATALYTIC CONSTANTS FOR THE BROMINATION OF ACETYLACETONE

Calculated values from $k_B = 0{\cdot}56\ (1/K_A)^{0{\cdot}52}$

Catalyst	k_B (obs.)	k_B (calc.)
CH_2ClCOO^-	36	34
CH_2OHCOO^-	94	108
CH_3COO^-	308	332
CMe_3COO^-	508	466
OH^-	$< 10^7$	9×10^7
H_2O	0·024	0·069

GENERAL DISCUSSION

The tables given for each ketone show that in each case the catalytic constants for the four carboxylate anions are well represented by an equation of the type $k_B = G(1/K_A)^\alpha$, so that the reactions studied are suitable ones for investigating the points mentioned in the introduction. In considering the variations in G and α we shall also include data obtained by K. J. Pedersen (1933, 1934*b*). For the bromination of acetoacetic ester he gives catalytic constants for the acetate, glycollate and monochloroacetate ions at 25°, while for the bromination of acetoacetic acid catalytic constants for the acetate and glycollate ions at 25° can be obtained by extrapolation from his data at 0° and 18°. The constants of the Brönsted equation have been recalculated using the dissociation constants employed in the present paper. Pedersen also gives data for the bromination of α-bromoacetoacetic ester and acid, but these are admittedly of lower accuracy and have not been included. The values for G and α are collected in the first two columns of table 18.

TABLE 18. CATALYSIS BY CARBOXYLATE ANIONS

Ketone	α	$\log_{10}G$	x	p	$\log_{10}R$
CH_3COCH_3	0·88	−9·04	3	2	−6·30
$CH_3COCH_2CH_2COCH_3$	0·89	−8·28	2	2	−5·16
CH_3COCH_2Cl	0·82	−6·19	2	1	−3·21
CH_3COCH_2Br	0·82	−5·93	2	1	−2·95
$CH_3COCHCl_2$	0·78	−5·14	1	1	−2·02
CH_3COCH_2COOEt	0·59	−1·34	—	1	+1·02
$CH_3COCH_2COCH_3$	0·52	−0·55	2	1	+1·83
CH_3COCH_2COOH	0·48	+0·65	—	1	+2·57

It will be seen that the exponent α changes steadily as the value of G changes through about ten powers of ten, thus confirming the suggestion that the extended Bronsted relation (equation (2)) is essentially a differential relation with a limited range of validity. The only other large change of exponent which has been recorded is in the acid catalysed rearrangement of three N-halogenacylanilides (Bell and Danckwerts 1939) and in this case the position is complicated by the fact that the reactions involve the migration of different halogen atoms.

The last column of table 18 contains the values of $\log_{10}R$, where R may be defined as the calculated catalytic constant for an acid of dissociation constant 10^{-4}, corrected for statistical differences between the different ketones. R is thus a measure of the ease with which a single hydrogen atom ionizes from the ketone in question. The first statistical correction depends on the fact that the initial ionization of the ketone is in some cases succeeded by the introduction of more than one halogen atom into the ketone. In order to convert rates of halogenation into rates of ionization it is therefore necessary to divide by a factor x, where x is the number of halogen atoms thus introduced. (In the case of acetoacetic ester and acid this correction is already allowed for in the method of calculation used.) The other statistical correction arises from the presence of more than one point in the ketone at which ionization can take place: if there are p equivalent points the ionization velocities must be divided by p to make them comparable. This kind of statistical correction has been previously applied to catalysts of different types, and in a number of cases it is difficult to decide on the correct value of p. Thus Bronsted (1928) considers that several hydrogen atoms bound to the same atom are kinetically independent, while Pedersen (1934*a*, 1938) counts them as one only. We have adopted the latter alternative: thus for acetone we take $p = 2$ (not $p = 6$), and for monochloroacetone $p = 1$ (not $p = 2$). The values of x and p are given in table 18, and R is then given by $\log_{10} R = \log_{10} G + 4\alpha - \log_{10} px$. It may be noted that a different choice of p would have little effect on the general correlation between R and α.

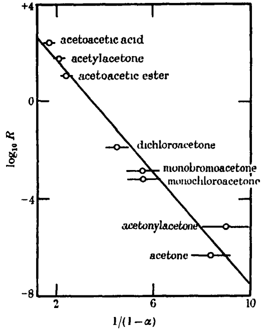


FIGURE 4

The observed increase of α with decreasing R cannot go on indefinitely, since α cannot become greater than unity. Figure 4 shows a plot of $\log_{10} R$ against $1/(1-\alpha)$ which will be shown in the next paper to represent the ratio of the slopes of the two potential energy curves at their point of intersection. The lengths of the horizontal lines indicate the errors corresponding to an error of ± 0·01 in α. The results are represented within experimental error by a linear relation between $1/(1-\alpha)$ and $\log_{10} R$.

The values observed for catalysis by hydroxyl ion and water are collected in table 19, together with the values calculated from the Brönsted equations valid for the carboxylate anions.

TABLE 19. CATALYSIS BY WATER MOLECULES AND HYDROXYL IONS

Ketone	k_{H_2O}		k_{OH^-}	
	Obs.	Calc.	Obs.	Calc.
CH_3COCH_3	5×10^{-10}	8×10^{-11}	$1{\cdot}5 \times 10$	$5{\cdot}2 \times 10^{4}$
$CH_3COCH_2CH_2COCH_3$	$2{\cdot}5 \times 10^{-9}$	$1{\cdot}5 \times 10^{-10}$	$1{\cdot}0 \times 10^{3}$	$5{\cdot}5 \times 10^{5}$
CH_3COCH_2Cl	$5{\cdot}8 \times 10^{-8}$	$2{\cdot}6 \times 10^{-8}$	$5{\cdot}6 \times 10^{3}$	$1{\cdot}5 \times 10^{6}$
CH_3COCH_2Br	$2{\cdot}9 \times 10^{-7}$	$4{\cdot}4 \times 10^{-8}$	$1{\cdot}2 \times 10^{4}$	$1{\cdot}0 \times 10^{7}$
$CH_3COCHCl_2$	$7{\cdot}9 \times 10^{-7}$	$3{\cdot}0 \times 10^{-7}$	$2{\cdot}7 \times 10^{4}$	$1{\cdot}4 \times 10^{7}$
CH_3COCH_2COOEt	$1{\cdot}3 \times 10^{-3}$	$4{\cdot}3 \times 10^{-3}$	—	9×10^{7}
$CH_3COCH_2COCH_3$	$1{\cdot}2 \times 10^{-2}$	$6{\cdot}9 \times 10^{-2}$	—	9×10^{7}
CH_3COCH_2COOH	$1{\cdot}4 \times 10^{-1}$	$6{\cdot}5 \times 10^{-1}$	-	$1{\cdot}6 \times 10^{8}$

It will be seen that the catalytic power of the hydroxyl ion is throughout several powers of ten smaller than the calculated value. This is qualitatively in accord with the behaviour found when the nature of the ketone is varied, i.e. a decrease in the exponent α when the velocity of the reaction is greatly increased. It is difficult to draw any quantitative conclusions since the figure taken to represent the basic strength of the hydroxyl ion

$$55{\cdot}5/K_w = 1/1{\cdot}8 \times 10^{-16}$$

can only do so approximately. The catalytic constants of the water molecule are throughout within a power of ten of the calculated values. However, little theoretical significance can be attached to this, since the water molecule is not of the same charge type as the anions on which the values of G and α are based, and the evaluations of the catalytic constant and the basic strength of water both involve the figure 55·5 as the supposed concentration of H_2O molecules in pure water.

The theoretical interpretation of the variation and absolute value of α is discussed in the next paper.

Our thanks are due to the Chemical Society for a grant for the purchase of materials.

REFERENCES

Bell, R. P. 1936 *Proc. Roy Soc.* A, **154**, 114

— 1938 *Trans Faraday Soc.* **34**, 229.

Bell, R. P. and Danckwerts, P. V. 1939 *J. Chem. Soc.* p. 1774.

Bell, R. P., Lidwell, O. M. and Vaughan-Jackson, M. W. 1936 *J. Chem. Soc.* p. 1794.

Bjerrum, N. and Unmack, A 1929 *K. danske vidensk. Selsk. Math. Fys. Medd.* **9**, no. 1.

Brönsted, J. N. 1928 *Chem. Rev.* **5**, 322.
Brönsted, J. N. and Pedersen, K. J. 1924 *Z. phys. Chem.* **108**, 185.
Dawson, H. M. and Carter, J. S. 1926 *J. Chem. Soc.* p. 2782.
Dawson, H. M. and Dean, N. C. 1926 *J. Chem. Soc.* p. 2872.
Dawson, H. M., Hall, G. V. and Key, A. 1928 *J. Chem. Soc.* p. 2844.
Dawson, H. M. and Hoskins, C. R. 1926 *Proc. Leeds Phil. Lit. Soc.* **1**, 108.
Dawson, H. M., Hoskins, C. R. and Smith, J. E. 1929 *J. Chem. Soc.* p. 1884.
Dawson, H. M. and Key, A. 1928 *J. Chem. Soc.* pp. 543, 1239, 1248.
Dawson, H. M. and Powis, F. 1913 *J. Chem. Soc* p. 2135
Dawson, H. M. and Smith, J. E. 1929 *J. Chem Soc* p. 2530.
— — 1930 *J. Chem. Soc.* p. 79.
Dawson, H. M. and Spivey, E. 1930 *J. Chem. Soc.* p. 2180.
Dippy, J. F. J. 1938 *J. Chem. Soc.* p. 1222.
Guggenheim, E. A. 1935 *Phil. Mag.* (7), **19**, 588.
Harned, H. S. and Ehlers, R. W. 1933 *J. Amer. Chem. Soc.* **35**, 652.
Horiuti, J. and Polanyi, M. 1935 *Acta Physicochim. U.R S S.* **2**, 505.
Levene, P. A. 1930 *Organic syntheses*, **10**, 12 New York.
Meyer, K. H. 1911 *Ber. dtsch. chem Ges.* **45**, 2843.
Nims, L. F. 1936 *J. Amer. Chem Soc.* **58**, 987.
Ogston, A G. 1936 *J. Chem Soc.* p. 1713.
Pedersen, K. J. 1933 *J. Phys. Chem* **37**, 751.
— 1934*a* *J. Phys. Chem.* **38**, 581.
— 1934*b* *J. Phys. Chem.* **38**, 99, 601.
— 1938 *Trans. Faraday Soc.* **34**, 237.
Puntambeker, S V. and Zoellner, E. A. 1932 *Organic syntheses*, Collected vol. **1**, pp. 451, 512. New York.
Shipsey, K. and Werner, E. A. 1913 *J. Chem. Soc.* p. 1255.
Tiessens, A. 1929 *Rec. Trav. chim. Pays-Bas*, **48**, 1066.
Watson, H. B. and Yates, E. D. 1932 *J. Chem. Soc.* p. 1207.
Wright, D. D. 1934 *J. Amer. Chem. Soc.* **56**, 314.

Potential energy curves in proton transfer reactions

BY R. P. BELL AND O. M. LIDWELL

(Communicated by C. N. Hinshelwood, F.R.S.—Received 23 April 1940)

The potential energy curves involved in proton transfer reactions are discussed, with special reference to the base-catalysed prototropy of ketones. The assumption of covalent binding throughout is inadequate to account in detail for the relations observed between basic strength and catalytic power. Approximate calculations from first principles show that the ionic state may well be of importance in this type of reaction, and the energy curves constructed on this basis give a much more satisfactory picture of the experimental facts.

INTRODUCTION

As pointed out in the introduction to the preceding paper (quoted as I), the rate determining step in reactions catalysed by acids or bases is a proton transfer reaction between the substrate and the catalyst. For basic catalysis (substrate SH, catalyst B) this may be written schematically

$$\mathrm{SH}+\mathrm{B}\rightarrow \mathrm{BH}^{+}+\mathrm{S}^{-}.$$

The velocity constants (k_B) of such reactions are related to the associated equilibrium constants by an equation of the form

$$d\log k_B = \alpha\{d\log K_B + d\log (K_A)_S\} = \alpha d\log K_R, \tag{1}$$

where K_B measures the basic strength of the catalyst, $(K_A)_S$ the acid strength of the substrate, and K_R is the equilibrium constant for the protolytic reaction between the substrate and the catalyst.* In the theoretical treatment of proton transfer reactions (Horiuti and Polanyi 1935; Bell 1936) the fraction α in equation (1) is related to the slopes of the potential energy curves for SH and BH at the point where they cross. if these slopes (both considered positive) are p and q respectively, then

$$\alpha = p/(p+q). \tag{2}$$

If the variations in k_B are not very large, the potential energy curves will be effectively linear over the relevant range, and α may be taken as a constant.

---

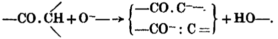
* The treatment given in this paper applies equally to catalysis by acids and by bases. For convenience we have considered basic catalysis throughout, but the corresponding application to acid catalysis should be obvious.

This is normally the case for catalysis of a given reaction by a series of related catalysts, and the resulting integrated equation is that originally proposed by Brönsted for expressing the relation between catalytic power and basic strength. Much greater variations in k_B can be obtained by varying the substrate, and the data given in the preceding paper (I) for the base catalysed prototropy of a series of ketones show that an increase of k_B by about nine powers of ten is accompanied by a steady decrease in α from 0·89 to 0·48. Further evidence for such variation in α comes from the catalytic constants observed for the hydroxyl ion in a number of reactions. For the reactions studied in (I), the decomposition of nitramide (Marlies and LaMer 1935), and the bromination of nitromethane (Pedersen 1932) the catalytic effect of the hydroxyl ion is smaller by several powers of ten than the value extrapolated from the catalytic effects of other anions. This again corresponds qualitatively to a decrease in α with increasing rate.

It is the purpose of the present paper to consider the variation and absolute value of α in relation to the nature of the potential energy curves involved in the reaction. In view of the experimental data outlined above we shall deal in the first instance with the prototropy of ketones catalysed by basic anions. The conclusions arrived at will however probably apply fairly generally to reactions catalysed by acids and bases.

COVALENT POTENTIAL ENERGY CURVES

The reaction involved in the transfer of a proton from a ketone to a carboxylate ion can be written



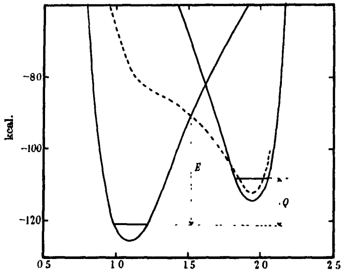
$$—\mathrm{CO}.\overset{\diagup}{\underset{\diagdown}{\mathrm{C}}}\mathrm{H} + \mathrm{O}^- \longrightarrow \left\{\begin{matrix} —\mathrm{CO}.\mathrm{C}^-\!— \\ —\mathrm{CO}^- : \mathrm{C}= \end{matrix}\right\} + \mathrm{HO}—. \qquad (3)$$

The ion formed from the ketone is considered as a mesomeric or resonance state between the two formulae given (Pedersen 1938). (The carboxylate ion probably ought also to be represented as a resonance state in which the two oxygen atoms are equivalent: however, no quantitative treatment of this state can be given, and we shall therefore treat the ion as containing a single oxygen atom with a negative charge.) The hydrogen atom is thus covalently bound in both the initial and the final states, of the system. It is therefore natural to investigate first how far the whole process can be interpreted in terms of covalent potential energy curves, it being assumed that these curves are not appreciably modified by the proximity of ions. This picture of the proton transfer has already been employed by Horiuti and Polanyi (1935). Since the process concerned involves no change in the number of ions, we shall also neglect the interaction of the solvent.

The curves for the links C—H and O—H can be assumed to be Morse curves, the form of which is given by the dissociation energies and the fundamental frequencies. We have taken for C—H $D_e = 104{\cdot}0$ kcal./mole, $\omega_e = 3160$ cm.$^{-1}$, and for O—H $D_e = 114{\cdot}8$ kcal./mole, $\omega_e = 3760$ cm.$^{-1}$. These curves are drawn in figure 1 (full lines), and the zero point energy levels indicated correspond to 4·4 kcal. for C—H and 5·3 kcal. for O—H. If the reaction is to take place at a measurable speed, the distance between the carbon and oxygen atoms during reaction must be smaller than 3·5 A, the normal distance between unbonded atoms (cf. Horiuti and Polanyi 1935). We have chosen a value of 2·9 A, which is little less than some crystallographic distances between unbonded carbon and oxygen atoms (cf. Moerman 1937) Taking the C—H distance as 1·08 A and the O—H distance as 0·96 A, this gives a distance of 0·86 A between the two equilibrium positions of the proton.

The energy difference between the two systems in equation (3) can be fixed approximately by the following considerations. The transfer of a proton from carbon to oxygen will undoubtedly be endothermic, since the ketones concerned are much weaker acids than the carboxylic acids. (The dissociation constant of the keto form of acetoacetic ester is about 10^{-11}; cf Pedersen 1938.) On the other hand, the difference in energy cannot be greater than the activation energy of the process, which is about 22 kcal. for acetone (Smith 1934) and about 13 kcal. for acetoacetic ester (Pedersen 1934). The relative position of the curves in figure 1 corresponds to an energy change Q of 10 kcal.

It will be seen from figure 1 that this picture of the proton transfer predicts an activation energy E of about 40 kcal., which is considerably greater than the observed values of 13–22 kcal. However, resonance between the two states will lower the energy near the intersection of the two curves, and it might also be lowered by a closer distance of approach of the two reactants, though any considerable reduction in this distance would introduce a large energy of repulsion. A more serious objection to the covalent picture arises from the slopes of the two curves at their points of intersection. Both the O—H and the C—H curves are practically straight lines over a range of 20–30 kcal. in the neighbourhood of their intersection, the former being the steeper by about 20 %. This means (cf. equation (2)) that the exponent α should have a value of about 0·45 and should be practically constant over a wide range of velocities. The observed values for a series of eight ketones (table 18, paper I) vary continuously between 0·48 and 0·89, while five of them are between 0·8 and 0·9, corresponding to a C—H curve 5–10 times as steep as the O—H curve. In this respect the covalent picture is quite inadequate to account for the facts. It may further be noted that since practically all proton transfer reactions take place between the links O—H, C—H and N—H, the covalent picture would lead us to expect that the value of α would be close to 0·5 for all reactions of this type, since the force constants and dissociation energies do not differ widely in these three bonds. Actually values as low as 0·3 and as high as 0·8 are frequently met with, so that it is very doubtful whether the covalent potential energy curves are in general sufficient to account for the observed behaviour.



distance from C nucleus, A

FIGURE 1

IONIC POTENTIAL ENERGY CURVES

Although there is no doubt that the hydrogen is covalently linked in isolated ketone or carboxylic acid molecules, it is possible that the ionic state may be of importance when the ketone and the carboxylate ion are close together, since the potential energy will be lowered by Coulomb interaction. The qualitative treatment previously given (Bell 1936) assumes that the binding is ionic, and an attempt at a quantitative treatment will be given in this section.

The free ion of the ketone is a resonance structure (equation (3)) in which the charge is probably concentrated chiefly on the oxygen atom. However, during the process of proton transfer the proton is near the carbon atom, and its positive charge will stabilize the negative charge on the carbon. The system to be studied can therefore be written $>\!\!-C^{-}\ldots H^{+}\ldots O^{-}-$. The external field of this system will vary very little with the position of the proton, so that solvent interaction can be neglected in calculating the shape of the curve. In calculating the potential due to the two ions we have assumed that each contains eight L electrons represented by hydrogen-like wave functions, and that the K electrons effectively form part of the nucleus. Calculations similar to those made by Pauling (1927*a*) then give for the potential energy Φ of a charge $+\epsilon$ distant R from the ion

$$\Phi = \epsilon^2\{\tfrac{1}{3}e^{-b}(b^3+6b^2+18b+24)-1\}/R, \tag{4}$$

where $b = (Z-S)\,R/a_0$, $a_0 = 0{\cdot}528$ A, Z = nuclear charge, and S = screening constant The screening constants used were obtained arbitrarily by multiplying the value for neon (4·26; Pauling 1927*b*) by $\frac{5}{8}$ for the carbon ion and by $\frac{7}{8}$ for the oxygen ion. The energy of the system $>\!\!-C^{-}\ldots H^{+}\ldots O^{-}-$ is then obtained by addition, including a term for the mutual energy of the two negative ions in which they are treated as point charges. No allowance has been made for polarization.

This treatment is clearly only approximate, but should serve to indicate the general shape of the potential energy curves. Figure 2 shows the calculated curves for different values of the carbon-oxygen internuclear distance d, the ordinates representing the distance of the proton from the carbon nucleus. The zero energy level in this diagram corresponds to infinite separation of the charges $>\!\!-C^{-}$, H^{+}, and $O^{-}-$ The asymptotic level of the O—H curve in figure 1 (taken as the energy zero) represents the state $>\!\!-C^{-}+H-O-$: hence the energy difference between these levels in the two diagrams is given by the energy of the process $-O^{-}+H^{+}$ (infinite separation)$\rightarrow -OH$. (Since in both cases the solvent interaction is being neglected when the reactants are close together the whole process can be considered to take place *in vacuo*.) We have used the following values.

$$-O + H \rightarrow OH + 115 \text{ kcal.},$$
$$H^{+} + \epsilon \rightarrow H + 311 \text{ kcal.},$$
$$-O + \epsilon \rightarrow -O^{-} + 64 \text{ kcal.}$$

The last value is in considerable doubt, published values for the electron affinity of oxygen ranging between 48 and 85 kcal. (cf. Goubeau, 1936, 1937). This gives

$$—O^- + H^+ \rightarrow —OH + 360 \text{ kcal.}$$

which makes the energy of the minimum in figure 2 (for a carbon-oxygen distance of 2·9 A) almost coincide with that of the right-hand minimum in figure 1. No significance can be attributed to this exact result, but the calculation shows that it is not unreasonable to suppose that the ionized state takes part in the reaction.

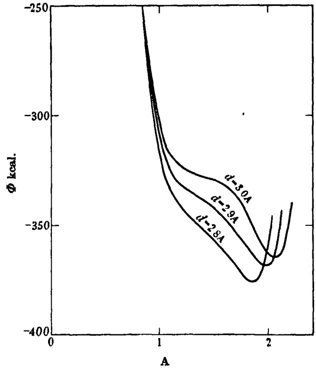


FIGURE 2

The broken curve in figure 1 represents the ionized state as calculated above. At the point where it intersects the covalent C—H curve the slopes of the two curves are very different, thus allowing values of α differing considerably from 0·5. The actual value of α corresponding to the curves shown in figure 1 is 0·85. Moreover, the general shape of the ionic curve is such as to allow considerable changes in α as the point of intersection moves along the curve, though the actual changes corresponding to the curve in figure 1 are considerably smaller than those found experimentally. However, it should be noted that the direct derivation of α in this way assumes that a change in the nature of the base will cause equal energy displacements in the covalent —OH curve and the ionic curve, while the method actually used for calculating the energy of the ionic state makes no allowance for any changes in the groups attached to the oxygen atom. Such an allowance could be made formally by introducing terms for dipole and polarizability effects, but in view of the approximate nature of the treatment a more detailed calculation is hardly justified. It does however seem probable that the difficulties which arise in a covalent picture of proton transfer are largely absent when the ionic state is taken into account.

The above general considerations should apply qualitatively to any protolytic reactions in which the base has a negative charge. If, however, the reaction is of the type $B_1H + B_2 \rightarrow B_1^- + B_2H^+$ (i.e. catalysis of an uncharged substrate by an uncharged acid or base) it is easily seen that the energy of the ionized state is much increased and is likely to be higher than that of the covalent state for all configurations We might therefore expect to find that for reactions of this type the exponent α will usually be in the neighbourhood of 0·5. There is some experimental support for this view, thus for catalysis by undissociated acid molecules α is 0·55 for the iodination of acetone (paper I), 0·4 for the depolymerization of dimeric dihydroxyacetone and glycollaldehyde (Bell and Baughan 1937; Bell and Hirst 1939), and 0·42 and 0·53 for the iodination of phosphorous and hypophosphorous acids respectively (Griffith, McKeown and Taylor 1940). On the other hand, the decomposition of nitramide catalysed by uncharged bases in three different solvents (Brönsted and Duus 1925, Brönsted and Vance 1933, Brönsted, Nicholson and Delbanco, 1934) gives an exponent of about 0·8, and it does not seem possible to account for this on the basis of covalent curves.

REFERENCES

Bell, R. P. 1936 *Proc. Roy Soc.* A, **154**, 114.

Bell, R. P. and Baughan, E. C 1937 *J. Chem. Soc.* p. 1947

Bell, R. P. and Hirst, J. P. H. 1939 *J. Chem. Soc.* p. 1777.

Brönsted, J. N. and Duus, H. C. 1925 *Z. phys. Chem.* **117**, 299.

Bronsted, J. N. and Vance, J. E. 1933 *Z. phys. Chem.* **163 A**, 240.

Bronsted, J. N., Nicholson, A. L and Delbanco, A. 1934 *Z. phys. Chem.* **169 A**, **379**.

Goubeau, J. 1936 *Z. phys. Chem.* **34 B**, 438.

Goubeau, J. and Klemm, W. 1937 *Z. phys. Chem.* **36 B**, 369.

Griffith, R. O, McKeown, A. and Taylor, R. P. 1940 *Trans. Faraday Soc.* **36**, **752**, **766**.

Horiuti, J. and Polanyi, M. 1935 *Acta Phys. U.R.S.S.* **2**, 505.
Marlies, C. A. and LaMer, V. K. 1935 *J. Amer. Chem. Soc.* **57**, 1812.
Moerman, N. F. 1937 *Rec. trav. chim. Pays-Bas*, **56**, 661.
Pauling, L. 1927*a* *Proc. Roy. Soc.* A, **114**, 181.
— 1927*b* *Z. Kristallogr.* **81**, 1.
Pedersen, K. J. 1932 *K. danske vidensk. Selsk.* **12**, 1.
— 1934 *J. Phys. Chem.* **38**, 99, 601.
— 1938 *Trans. Faraday Soc.* **34**, 237.
Smith, G. F. 1934 *J. Chem. Soc.* p. 1744.

General bi-harmonic analysis for a plate containing circular holes

By A. E. Green, Ph.D.

(*Communicated by G. I. Taylor, F.R.S.—Received* 24 *April* 1940)

A general solution is given for problems of generalized plane stress distributions in an infinite plate which contains circular holes of varying sizes in any positions, subject only to certain conditions of convergence of the solution. The method for extending the results to allow for the effect of one or two straight boundaries is indicated.

As a particular case of the general solution the problem of the stress distribution in an infinite plate under tension containing three holes in a row, is discussed, and a few numerical results are given. These results are compared with experiments which were carried out by Mr P. L. Capper. The comparison is incomplete as the experiments were done for a finite plate and the influence of the edges of the plate on the numerical values of the stresses is considerable. Agreement, however, is found for the general character of some of the stresses.

INTRODUCTION

1. The distribution of stress in plates containing holes has been studied by a number of writers (see references). The solutions which have been obtained have been confined to a group of problems in which the boundaries are a set of equal circles together with, in some cases, a pair of parallel straight lines. As is pointed out by Howland and Knight (1939) no general method of solution is given and the circles, their number and relative positions, are restricted by an invariancy condition which demands that the circles and their boundary conditions remain invariant under one or more of a group of transformations.

Numerical results have been obtained for the stress distribution in a plate under tension containing one circular hole and these results have been compared with experiment (Jeffery 1920; Howland 1930; Howland and Stevenson 1933) Numerical work has also been carried out for a plate under tension containing an infinite row of equal circular holes (Howland 1935) and for various problems connected with two circular holes in an infinite plate (Jeffery 1920) No numerical results appear to have been given for any of the other problems which have been considered by Howland and Knight, and so far as the present writer is aware, no experimental results exist for these problems

Experimental results do, however, exist for stresses in a strip under tension when the strip contains a group of three holes or a group of six holes (Coker and Filon 1931, chap. VI), but these stresses cannot be evaluated by the methods which have been used by previous writers. It would thus seem that the solution of these problems should be included in any further work on the distribution of stress in plates containing holes. It has been found that it is possible to give a general method of solution which is applicable for any problem of stress distribution in infinite plates, plates bounded by one straight edge, or plates bounded by two straight edges, which contain any number of circular holes of various sizes distributed in any manner, subject only to certain conditions of convergence of the solution. Previous solutions may be deduced as special cases but in many problems it is best to use from the outset any properties of symmetry that may exist as this greatly simplifies the work.

In the present paper a general solution is given for problems of stress distributions in an infinite plate which contains holes of varying sizes. The method may easily be extended to allow for the effect of straight boundaries. This extension is only outlined here because the algebra required for the general case is very complicated. Also it would appear to have little value as, owing to the very heavy nature of the arithmetic, it seems unlikely that the solution would be completed by the necessary numerical work.

As a particular case of the general solution the problem of the stress distribution in an infinite plate under tension containing three circular holes in a row, is discussed, and a few numerical results are given. These results are compared with experiments which were carried out by Mr P. L. Capper, but the comparison is incomplete as the experiments were done for a finite plate and the influence of the edges of the plate is considerable. It has, however, already been indicated that the arithmetical work which is required for a finite plate seems to be out of all proportion to the value of any results that would be obtained.

When the present work was started it was intended that other special cases would be considered but these cases have now been adequately dealt with by special methods in a recent paper by Howland and Knight (1939).

THE GENERAL METHOD OF SOLUTION

2. Consider a uniform plate of infinite extent containing circular holes of varying radii. It is supposed to be in a state of generalized plane stress defined by a stress function χ, where

$$\nabla_1^4 \chi = 0. \tag{2·1}$$

The boundary conditions to be satisfied are.

(*a*) the stresses $\widehat{xx} \equiv \dfrac{\partial^2 \chi}{\partial y^2}, \quad \widehat{yy} \equiv \dfrac{\partial^2 \chi}{\partial x^2}, \quad \widehat{xy} \equiv -\dfrac{\partial^2 \chi}{\partial x \, \partial y},$

must tend to constant values at infinity;

(*b*) on the circumference of each hole, the stresses

$$\widehat{rr} \equiv \frac{1}{r^2}\frac{\partial^2 \chi}{\partial \theta^2} + \frac{1}{r}\frac{\partial \chi}{\partial r} = 0, \quad \widehat{r\theta} \equiv -\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial \chi}{\partial \theta}\right) = 0.$$

At the centre of each hole a set of rectangular axes is taken parallel to fixed directions. For example, if O_r is the centre of the rth hole, axes are taken at O_r which may be denoted by $O_r x_r$, $O_r y_r$. Referred to these axes at O_r a complex co-ordinate z_r may be defined as

$$z_r = x_r + i y_r = R_r (\sin \theta_r + i \cos \theta_r), \tag{2·2}$$

where (R_r, θ_r) are plane polar co-ordinates having their pole at O_r. The angle θ_r is measured from the axis of y_r as this is the most convenient notation to use if the analysis has to be extended in order to consider the effect of plane boundaries.

It is convenient to introduce non-dimensional co-ordinates as follows:

$$\zeta_r = \xi_r + i\eta_r = \rho_r(\sin \theta_r + i \cos \theta_r) = i\rho_r e^{-i\theta_r}, \tag{2·3}$$

where $\quad \zeta_r = z_r/h, \quad \xi_r = x_r/h, \quad \eta_r = y_r/h, \quad \rho_r = R_r/h,$

and h is a typical length which may, for example, be the minimum distance between the centres of the circles. The relations between the co-ordinate axes at the centres O_r, O_s of two typical circles may be expressed in the form

$$\zeta_r = \zeta_s + i h_{rs} e^{-i\alpha_{rs}}, \tag{2·4}$$

so that $h_{rs} = O_r O_s / h$ and α_{rs} is the angle between $O_r O_s$ and the axis $O_r y_r$ at O_r.

If there were no holes in the plate, then the stresses produced in it when it is acted on by a system of generalized plane stress may be derived from a stress function χ_0. In order to allow for the effect of the holes a fundamental set of stress functions is required for each hole which give zero stresses at infinity and which give single-valued expressions for the stresses and displacements. The required functions are

$$\log \rho_r,\ \theta_r;\quad \rho_r^{-n}\cos n\theta_r,\ \rho_r^{-n}\sin n\theta_r,\quad (n = 1, 2, \ldots)$$

$$\rho_r^{-n}\cos (n+2)\,\theta_r,\ \rho_r^{-n}\sin (n+2)\,\theta_r,\quad (n = 0, 1, 2, \ldots)$$

where r takes all integral values from 1 to m corresponding to the m holes. The general stress function for the problem is then

$$\begin{aligned}\chi = \chi_0 &+ \sum_{s=1}^{m}\{{}^sA_0 \log \rho_s + {}^sB_0\theta_s\}\\ &+ \sum_{s=1}^{m}\sum_{n=1}^{\infty}\{{}^sA_n\rho_s^{-n}\cos n\theta_s + {}^sB_n\rho_s^{-n}\sin n\theta_s\}\\ &+ \sum_{s=1}^{m}\sum_{n=0}^{\infty}\{{}^sC_n\rho_s^{-n}\cos (n+2)\,\theta_s + {}^sD_n\rho_s^{-n}\sin (n+2)\,\theta_s\},\end{aligned} \tag{2·5}$$

where ${}^sA_n$, ${}^sB_n$, ${}^sC_n$, ${}^sD_n$ are constants which are determined by the boundary conditions at the circumference of each hole. The stress function may be taken in the form (2·5) which is independent of the elastic constants provided that the force resultants of the tractions applied to each hole separately vanish. In order to evaluate these constants it is necessary to express the stress function χ in terms of the plane polar co-ordinates ρ_r, θ_r, and this work is carried out in the next section.

EXPANSIONS ABOUT THE CENTRE OF ANY CIRCLE

3. By using (2·3) and (2·4) the fundamental stress functions which are connected with the origin O_s can be expressed in terms of ρ_r and θ_r. The use of the complex variable greatly simplifies the work and the final results are as follows:

$$\left.\begin{aligned}\log \rho_s &= \sum_{k=1}^{\infty}\rho_r^k({}^0a_k^{s,r}\cos k\theta_r + {}^0b_k^{s,r}\sin k\theta_r),\\ \theta_s &= \sum_{k=1}^{\infty}\rho_r^k({}^0c_k^{s,r}\cos k\theta_r + {}^0d_k^{s,r}\sin k\theta_r),\end{aligned}\right\} \tag{3·1}$$

$$\left.\begin{aligned}\rho_s^{-n}\cos n\theta_s &= \sum_{k=1}^{\infty}\rho_r^k({}^na_k^{s,r}\cos k\theta_r + {}^nb_k^{s,r}\sin k\theta_r),\\ \rho_s^{-n}\sin n\theta_s &= \sum_{k=1}^{\infty}\rho_r^k({}^nc_k^{s,r}\cos k\theta_r + {}^nd_k^{s,r}\sin k\theta_r),\end{aligned}\right\} \tag{3·2}$$

where $n = 1, 2, 3, \ldots$ and

$$\left.\begin{aligned} \rho_s^{-n}\cos(n+2)\,\theta_s &= \sum_{k=0}^{\infty}\rho_r^k\{{}^ne_k^{s,r}\cos k\theta_r + {}^nf_k^{s,r}\sin k\theta_r \\ &\qquad + \rho_r^2({}^ng_k^{s,r}\cos k\theta_r + {}^nh_k^{s,r}\sin k\theta_r)\}, \\ \rho_s^{-n}\sin(n+2)\,\theta_s &= \sum_{k=0}^{\infty}\rho_r^k\{{}^nu_k^{s,r}\cos k\theta_r + {}^nv_k^{s,r}\sin k\theta_r \\ &\qquad + \rho_r^2({}^nx_k^{s,r}\cos k\theta_r + {}^ny_k^{s,r}\sin k\theta_r)\}, \end{aligned}\right\} \tag{3·3}$$

where $n = 0, 1, 2, 3, \ldots$ In the above expansions some of the constant terms are omitted as they contribute nothing to the stresses. The coefficients in the expansions are

$$\left.\begin{aligned} {}^0a_k^{s,r} = {}^0d_k^{s,r} &= -\frac{1}{kh_{rs}^k}\cos k\alpha_{rs}, \\ {}^0b_k^{s,r} = -{}^0c_k^{s,r} &= -\frac{1}{kh_{rs}^k}\sin k\alpha_{rs}, \end{aligned}\right| \quad (k \geqslant 1) \tag{3·4}$$

$$\left.\begin{aligned} {}^na_k^{s,r} = -{}^nd_k^{s,r} &= \frac{(-)^n}{h_{rs}^{n+k}}\binom{k+n-1}{k}\cos(k+n)\,\alpha_{rs}, \\ {}^nb_k^{s,r} = {}^nc_k^{s,r} &= \frac{(-)^n}{h_{rs}^{n+k}}\binom{k+n-1}{k}\sin(k+n)\,\alpha_{rs}. \end{aligned}\right\} \quad (n \geqslant 1,\ k \geqslant 1) \tag{3·5}$$

Also

$$\left.\begin{aligned} {}^ne_1^{s,r} &= \frac{(-)^n(n+1)}{h_{rs}^{n+1}}\cos(n+3)\,\alpha_{rs} - \frac{(-)^n}{h_{rs}^{n+1}}\cos(n+1)\,\alpha_{rs}, \\ {}^nf_1^{s,r} &= \frac{(-)^n(n+1)}{h_{rs}^{n+1}}\sin(n+3)\,\alpha_{rs} + \frac{(-)^n}{h_{rs}^{n+1}}\sin(n+1)\,\alpha_{rs}, \\ {}^nu_1^{s,r} &= \frac{(-)^n(n+1)}{h_{rs}^{n+1}}\sin(n+3)\,\alpha_{rs} - \frac{(-)^n}{h_{rs}^{n+1}}\sin(n+1)\,\alpha_{rs}, \\ {}^nv_1^{s,r} &= -\frac{(-)^n(n+1)}{h_{rs}^{n+1}}\cos(n+3)\,\alpha_{rs} - \frac{(-)^n}{h_{rs}^{n+1}}\cos(n+1)\,\alpha_{rs}, \end{aligned}\right\} \quad (n \geqslant 0) \tag{3·6}$$

$$\left.\begin{aligned} {}^ne_k^{s,r} = -{}^nv_k^{s,r} &= \frac{(-)^n}{h_{rs}^{n+k}}\binom{n+k}{k}\cos(k+n+2)\,\alpha_{rs}, \\ {}^nf_k^{s,r} = {}^nu_k^{s,r} &= \frac{(-)^n}{h_{rs}^{n+k}}\binom{n+k}{k}\sin(k+n+2)\,\alpha_{rs}, \end{aligned}\right\} \quad (n \geqslant 0,\ k \geqslant 2) \tag{3·7}$$

$$\left.\begin{aligned} {}^ng_k^{s,r} = -{}^ny_k^{s,r} &= \frac{(-)^{n+1}}{h_{rs}^{n+k+2}}\binom{n+k+1}{k+1}\cos(k+n+2)\,\alpha_{rs}, \\ {}^nh_k^{s,r} = {}^nx_k^{s,r} &= \frac{(-)^{n+1}}{h_{rs}^{n+k+2}}\binom{n+k+1}{k+1}\sin(k+n+2)\,\alpha_{rs}. \end{aligned}\right\} \quad (n \geqslant 0,\ k \geqslant 0) \tag{3·8}$$

The above expansions converge if $\rho_r < h_{rs}$.

EVALUATION OF THE CONSTANTS

4. The stresses corresponding to χ_0 can be expressed in terms of ρ_r and θ_r, and at the circumference of the circle $\rho_r = \lambda_r$ they take the forms

$$\widehat{rr} = \sum_{k=1}^{\infty} {}^rU_k \sin k\theta_r + \sum_{k=0}^{\infty} {}^rS_k \cos k\theta_r,\ \widehat{r\theta} = \sum_{k=1}^{\infty} {}^rT_k \sin k\theta_r + \sum_{k=0}^{\infty} {}^rV_k \cos k\theta_r. \tag{4·1}$$

With the help of equations (3·1)–(3·3) the stress function χ (2·5) may be expressed in terms of ρ_r and θ_r and the corresponding values of $\widehat{rr}$ and $\widehat{r\theta}$ are then easily obtained as series of sines and cosines of θ_r whose coefficients are functions of ρ_r. In order that $\widehat{rr}$ and $\widehat{r\theta}$ may be zero when $\rho_r = \lambda_r$ for all values of θ_r the coefficients of the sines and cosines must be zero. These boundary conditions give the following sets of equations:

$${}^rS_0 h^2\lambda_r^2 + {}^rA_0 + 2\sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \lambda_r^2({}^sC_k\, {}^kg_0^{s,r} + {}^sD_k\, {}^kx_0^{s,r}) = 0, \tag{4·2}$$

$$\left.\begin{aligned} \frac{{}^rS_1 h^2}{2} - \frac{{}^rA_1}{\lambda_r^3} + \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \lambda_r({}^sC_k\, {}^kg_1^{s,r} + {}^sD_k\, {}^kx_1^{s,r}) = 0, \\ \frac{{}^rT_1 h^2}{2} - \frac{{}^rA_1}{\lambda_r^3} + \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \lambda_r({}^sC_k\, {}^kg_1^{s,r} + {}^sD_k\, {}^kx_1^{s,r}) = 0, \end{aligned}\right\} \tag{4·3}$$

and for $n \geqslant 2$,

$$\begin{aligned} &{}^rS_n h^2 - \frac{n(n+1)}{\lambda_r^{n+2}}\,{}^rA_n - \frac{(n-1)(n+2)}{\lambda_r^n}\,{}^rC_{n-2} \\ &\quad - \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \{(n-1)\,n\,\lambda_r^{n-2}({}^sA_k\, {}^ka_n^{s,r} + {}^sB_k\, {}^kc_n^{s,r}) \\ &\qquad + {}^sC_k\, \lambda_r^{n-2}[(n-1)\,n\, {}^ke_n^{s,r} + (n-2)(n+1)\, {}^kg_n^{s,r}\, \lambda_r^2] \\ &\qquad + {}^sD_k\, \lambda_r^{n-2}[(n-1)\,n\, {}^ku_n^{s,r} + (n-2)(n+1)\, {}^kx_n^{s,r}\, \lambda_r^2]\} = 0, \end{aligned} \tag{4·4}$$

$$\begin{aligned} &{}^rT_n h^2 - \frac{n(n+1)}{\lambda_r^{n+2}}\,{}^rA_n - \frac{(n-1)\,n}{\lambda_r^n}\,{}^rC_{n-2} \\ &\quad + \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \{(n-1)\,n\,\lambda_r^{n-2}({}^sA_k\, {}^ka_n^{s,r} + {}^sB_k\, {}^kc_n^{s,r}) \\ &\qquad + {}^sC_k\, \lambda_r^{n-2}[(n-1)\,n\, {}^ke_n^{s,r} + n(n+1)\, {}^kg_n^{s,r}\, \lambda_r^2] \\ &\qquad + {}^sD_k\, \lambda_r^{n-2}[(n-1)\,n\, {}^ku_n^{s,r} + n(n+1)\, {}^kx_n^{s,r}\, \lambda_r^2]\} = 0. \end{aligned} \tag{4·5}$$

Also
$$ {}^{r}V_0 h^2 \lambda_r^2 + {}^{r}B_0 = 0, \tag{4·6}$$

$$\left.\begin{aligned} \frac{{}^{r}U_1 h^2}{2} - \frac{{}^{r}B_1}{\lambda_r^3} + \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \lambda_r ({}^{s}C_k\, {}^{k}h_1^{s,r} + {}^{s}D_k\, {}^{k}y_1^{s,r}) = 0, \\ \frac{{}^{r}V_1 h^2}{2} + \frac{{}^{r}B_1}{\lambda_r^3} - \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \lambda_r ({}^{s}C_k\, {}^{k}h_1^{s,r} + {}^{s}D_k\, {}^{k}y_1^{s,r}) = 0, \end{aligned}\right\} \tag{4·7}$$

and for $n \geqslant 2$,

$$\begin{aligned} {}^{r}U_n h^2 - \frac{n(n+1)}{\lambda_r^{n+2}} {}^{r}B_n - \frac{(n-1)(n+2)}{\lambda_r^n} {}^{r}D_{n-2} \\ - \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \{(n-1)\, n\, \lambda_r^{n-2} ({}^{s}A_k\, {}^{k}b_n^{s,r} + {}^{s}B_k\, {}^{k}d_n^{s,r}) \\ + {}^{s}C_k \lambda_r^{n-2} [(n-1)\, n\, {}^{k}f_n^{s,r} + (n-2)(n+1)\, {}^{k}h_n^{s,r} \lambda_r^2] \\ + {}^{s}D_k \lambda_r^{n-2} [(n-1)\, n\, {}^{k}v_n^{s,r} + (n-2)(n+1)\, {}^{k}y_n^{s,r} \lambda_r^2]\} = 0, \end{aligned} \tag{4·8}$$

$$\begin{aligned} {}^{r}V_n h^2 + \frac{n(n+1)}{\lambda_r^{n+2}} {}^{r}B_n + \frac{(n-1)\, n}{\lambda_r^n} {}^{r}D_{n-2} \\ - \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \{(n-1)\, n\, \lambda_r^{n-2} ({}^{s}A_k\, {}^{k}b_n^{s,r} + {}^{s}B_k\, {}^{k}d_n^{s,r}) \\ + {}^{s}C_k \lambda_r^{n-2} [(n-1)\, n\, {}^{k}f_n^{s,r} + n(n+1)\, {}^{k}h_n^{s,r} \lambda_r^2] \\ + {}^{s}D_k \lambda_r^{n-2} [(n-1)\, n\, {}^{k}v_n^{s,r} + n(n+1)\, {}^{k}y_n^{s,r} \lambda_r^2]\} = 0. \end{aligned} \tag{4·9}$$

The symbol $\sum_{s=1}^{m}{}'$ means the sum for all values of s from 1 to m, excluding the value $s = r$. The above equations were derived from the boundary conditions at the edge of a typical hole whose centre is at O_r. Similar sets of equations may be obtained from the boundary conditions at the edges of all the other holes and so r will take all values from 1 to m.

The equations (4·3) and (4·7) give ${}^{r}S_1 = {}^{r}T_1$ and ${}^{r}U_1 = -{}^{r}V_1$ which are the conditions that the stresses derived from χ_0 should produce zero force resultants on each hole.

The equations (4·2)–(4·9) can be replaced by the equations

$$ -{}^{r}S_0 h^2 \lambda_r^2 = {}^{r}A_0 - \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \{{}^{s}C_k\, {}^{k}\gamma_0^{s,r} + {}^{s}D_k\, {}^{k}\delta_0^{s,r}\}, \tag{4·10}$$

$$\begin{aligned} \frac{h^2\{(n+2)\, {}^{r}T_n - n\, {}^{r}S_n\} \lambda_r^{n+2}}{2n(n+1)} = {}^{r}A_n - \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \{{}^{s}A_k\, {}^{k}\alpha_n^{s,r} + {}^{s}B_k\, {}^{k}\beta_n^{s,r} \\ + {}^{s}C_k\, {}^{k}\gamma_n^{s,r} + {}^{s}D_k\, {}^{k}\delta_n^{s,r}\}, \end{aligned} \tag{4·11}$$

$$\begin{aligned} \frac{h^2({}^{r}S_{n+2} - {}^{r}T_{n+2}) \lambda_r^{n+2}}{2(n+1)} = {}^{r}C_n - \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \{{}^{s}A_k\, {}^{k}\kappa_{n+2}^{s,r} + {}^{s}B_k\, {}^{k}\mu_{n+2}^{s,r} \\ + {}^{s}C_k\, {}^{k}\nu_{n+2}^{s,r} + {}^{s}D_k\, {}^{k}\omega_{n+2}^{s,r}\}. \end{aligned} \tag{4·12}$$

Also
$$-{}^rV_0 h^2\lambda_r^2 = {}^rB_0, \tag{4·13}$$

$$-\frac{h^2\{(n+2)\,{}^rV_n + n\,{}^rU_n\}\,\lambda_r^{n+2}}{2n(n+1)} = {}^rB_n - \sum_{k=0}^{\infty}\sum_{s=1}^{m}{}' \{{}^sA_k\,{}^k\xi_n^{s,r} + {}^sB_k\,{}^k\eta_n^{s,r} + {}^sC_k\,{}^k\zeta_n^{s,r} + {}^sD_k\,{}^k\tau_n^{s,r}\}, \tag{4·14}$$

$$\frac{h^2({}^rU_{n+2} + {}^rV_{n+2})\,\lambda_r^{n+2}}{2(n+1)} = {}^rD_n - \sum_{k=0}^{\infty}\sum_{s=1}^{m}{}' \{{}^sA_k\,{}^k\epsilon_{n+2}^{s,r} + {}^sB_k\,{}^k\theta_{n+2}^{s,r} + {}^sC_k\,{}^k\phi_{n+2}^{s,r} + {}^sD_k\,{}^k\psi_{n+2}^{s,r}\}. \tag{4·15}$$

The equations (4·11) and (4·14) hold for all values of r from 1 to m and for $n \geqslant 1$ if it is remembered that ${}^rS_1 = {}^rT_1$ and ${}^rU_1 = -{}^rV_1$ The equations (4·12) and (4·15) are valid for all values of r from 1 to m and for $n \geqslant 0$. The new coefficients in the above equations are

$$\left.\begin{aligned} {}^k\alpha_n^{s,r} &= (n-1)\,{}^ka_n^{s,r}\,\lambda_r^{2n}, \\ {}^k\beta_n^{s,r} &= (n-1)\,{}^kc_n^{s,r}\,\lambda_r^{2n}, \\ {}^k\gamma_n^{s,r} &= \lambda_r^{2n}\{(n-1)\,{}^ke_n^{s,r} + n\,{}^kg_n^{s,r}\,\lambda_r^2\}, \\ {}^k\delta_n^{s,r} &= \lambda_r^{2n}\{(n-1)\,{}^ku_n^{s,r} + n\,{}^kx_n^{s,r}\,\lambda_r^2\}, \end{aligned}\right\} \tag{4·16}$$

$$\left.\begin{aligned} {}^k\xi_n^{s,r} &= (n-1)\,{}^kb_n^{s,r}\,\lambda_r^{2n}, \\ {}^k\eta_n^{s,r} &= (n-1)\,{}^kd_n^{s,r}\,\lambda_r^{2n}, \\ {}^k\zeta_n^{s,r} &= \lambda_r^{2n}\{(n-1)\,{}^kf_n^{s,r} + n\,{}^kh_n^{s,r}\,\lambda_r^2\}, \\ {}^k\tau_n^{s,r} &= \lambda_r^{2n}\{(n-1)\,{}^kv_n^{s,r} + n\,{}^ky_n^{s,r}\,\lambda_r^2\}, \end{aligned}\right\} \tag{4·17}$$

for $k \geqslant 0$, $n \geqslant 1$ and all values of s from 1 to m, excluding $s = r$;

$$\left.\begin{aligned} {}^k\kappa_n^{s,r} &= -n\,{}^ka_n^{s,r}\,\lambda_r^{2n-2}, \\ {}^k\mu_n^{s,r} &= -n\,{}^kc_n^{s,r}\,\lambda_r^{2n-2}, \\ {}^k\nu_n^{s,r} &= -\lambda_r^{2n-2}\{n\,{}^ke_n^{s,r} + (n+1)\,{}^kg_n^{s,r}\,\lambda_r^2\}, \\ {}^k\omega_n^{s,r} &= -\lambda_r^{2n-2}\{n\,{}^ku_n^{s,r} + (n+1)\,{}^kx_n^{s,r}\,\lambda_r^2\}, \end{aligned}\right\} \tag{4·18}$$

$$\left.\begin{aligned} {}^k\epsilon_n^{s,r} &= -n\,{}^kb_n^{s,r}\,\lambda_r^{2n-2}, \\ {}^k\theta_n^{s,r} &= -n\,{}^kd_n^{s,r}\,\lambda_r^{2n-2}, \\ {}^k\phi_n^{s,r} &= -\lambda_r^{2n-2}\{n\,{}^kf_n^{s,r} + (n+1)\,{}^kh_n^{s,r}\,\lambda_r^2\}, \\ {}^k\psi_n^{s,r} &= -\lambda_r^{2n-2}\{n\,{}^kv_n^{s,r} + (n+1)\,{}^ky_n^{s,r}\,\lambda_r^2\}, \end{aligned}\right\} \tag{4·19}$$

for $k \geqslant 0$, $n \geqslant 2$ and all values of s from 1 to m excluding $s = r$.

When $n = 0$, $k \geqslant 0$,

$${}^k\gamma_0^{s,r} = -{}^kg_0^{s,r}\,2\lambda_r^2, \quad {}^k\delta_0^{s,r} = -{}^kx_0^{s,r}\,2\lambda_r^2. \tag{4·20}$$

all the remaining coefficients being zero.

A formal solution of equations (4·10)–(4·15) is

$$\left.\begin{aligned} {}^{r}A_n &= \sum_{p=0}^{\infty} {}^{r}A_n^{(p)}, & {}^{r}B_n &= \sum_{p=0}^{\infty} {}^{r}B_n^{(p)}, \\ {}^{r}C_n &= \sum_{p=0}^{\infty} {}^{r}C_n^{(p)}, & {}^{r}D_n &= \sum_{p=0}^{\infty} {}^{r}D_n^{(p)}, \end{aligned}\right\} \tag{4·21}$$

where

$${}^{r}A_0^{(0)} = -{}^{r}S_0 h^2 \lambda_r^2, \tag{4·22}$$

$${}^{r}A_n^{(0)} = \frac{h^2\{(n+2)\,{}^{r}T_n - n\,{}^{r}S_n\}\,\lambda_r^{n+2}}{2n(n+1)}, \quad (n \geqslant 1) \tag{4·23}$$

$${}^{r}C_n^{(0)} = \frac{h^2({}^{r}S_{n+2} - {}^{r}T_{n+2})\,\lambda_r^{n+2}}{2(n+1)}, \quad (n \geqslant 0) \tag{4·24}$$

$${}^{r}B_n^{(0)} = -\frac{h^2\{(n+2)\,{}^{r}V_n + n\,{}^{r}U_n\}\,\lambda_r^{n+2}}{2n(n+1)}, \quad (n \geqslant 1) \tag{4·25}$$

$${}^{r}D_n^{(0)} = \frac{h^2({}^{r}U_{n+2} + {}^{r}V_{n+2})\,\lambda_r^{n+2}}{2(n+1)}; \quad (n \geqslant 0) \tag{4·26}$$

while, for $p \geqslant 0$,

$${}^{r}A_n^{(p+1)} = \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \{{}^{s}A_k^{(p)}\,{}^{k}\alpha_n^{s,r} + {}^{s}B_k^{(p)}\,{}^{k}\beta_n^{s,r} + {}^{s}C_k^{(p)}\,{}^{k}\gamma_n^{s,r} + {}^{s}D_k^{(p)}\,{}^{k}\delta_n^{s,r}\}, \tag{4·27}$$

$${}^{r}B_n^{(p+1)} = \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \{{}^{s}A_k^{(p)}\,{}^{k}\xi_n^{s,r} + {}^{s}B_k^{(p)}\,{}^{k}\eta_n^{s,r} + {}^{s}C_k^{(p)}\,{}^{k}\zeta_n^{s,r} + {}^{s}D_k^{(p)}\,{}^{k}\tau_n^{s,r}\}, \tag{4·28}$$

$${}^{r}C_n^{(p+1)} = \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \{{}^{s}A_k^{(p)}\,{}^{k}\kappa_{n+2}^{s,r} + {}^{s}B_k^{(p)}\,{}^{k}\mu_{n+2}^{s,r} + {}^{s}C_k^{(p)}\,{}^{k}\nu_{n+2}^{s,r} + {}^{s}D_k^{(p)}\,{}^{k}\omega_{n+2}^{s,r}\}, \tag{4·29}$$

$${}^{r}D_n^{(p+1)} = \sum_{k=0}^{\infty} \sum_{s=1}^{m}{}' \{{}^{s}A_k^{(p)}\,{}^{k}\epsilon_{n+2}^{s,r} + {}^{s}B_k^{(p)}\,{}^{k}\theta_{n+2}^{s,r} + {}^{s}C_k^{(p)}\,{}^{k}\phi_{n+2}^{s,r} + {}^{s}D_k^{(p)}\,{}^{k}\psi_{n+2}^{s,r}\}, \tag{4·30}$$

for all values of n and for all values of r from 1 to m.

The validity of this solution is established if the series (4·27)–(4·30) are convergent. Suitable inequalities for ${}^{k}\alpha_n^{s,r}$ etc. which are required for a proof of convergence may easily be found but it does not seem possible to obtain any really useful general conditions of convergence when the actual positions of the circles are unknown. The problem of convergence is best dealt with separately in each particular case which is considered.

THE EFFECT OF STRAIGHT BOUNDARIES

5. Suppose now that the plate is bounded by two parallel straight lines and that the m sets of non-dimensional axes (ξ_r, η_r) are chosen so that, referred to the rth set at O_r, the lines are defined by $\eta_r = 1 \pm \alpha_r$, $(-1 \leqslant \alpha_r \leqslant 1)$. The typical length h has been taken to be half the breadth of the strip and $h\alpha_r$ is the distance of O_r from the centre line of the strip.

The fundamental stress functions which were used in § 2 have now to be modified to allow for the effect of the straight boundaries. Thus, if χ_r represents any of the fundamental functions for the origin O_r a suitable function χ_r' may be added to χ_r so that $\chi_r + \chi_r'$ gives zero stresses at infinity and zero normal and shear stresses on the straight boundaries. The functions χ_r' are found by a method which was used by Howland (1929) but the details of the work are omitted here. The fundamental stress functions for the origin O_r are found to be

$$\log\rho_r + \int_0^\infty \{e^{u\alpha_r} + e^{-u\alpha_r}\}\{u\mathrm{C}' - u(\eta_r+\alpha_r)\,\mathrm{S}' + e^{-u}s\mathrm{C}'\}\frac{\cos u\xi}{u\Sigma}\,du$$
$$+ \int_0^\infty \{e^{u\alpha_r} - e^{-u\alpha_r}\}\{u\mathrm{S}' - u(\eta_r+\alpha_r)\,\mathrm{C}' + e^{-u}c\mathrm{S}'\}\frac{\cos u\xi}{u\Sigma'}\,du, \tag{5·1}$$

$$\theta_r - \int_0^\infty \{e^{u\alpha_r} - e^{-u\alpha_r}\}\{u\mathrm{C}' - u(\eta_r+\alpha_r)\,\mathrm{S}' + e^{-u}s\mathrm{C}'\}\frac{\sin u\xi}{u\Sigma}\,du$$
$$- \int_0^\infty \{e^{u\alpha_r} + e^{-u\alpha_r}\}\{u\mathrm{S}' - u(\eta_r+\alpha_r)\,\mathrm{C}' + e^{-u}c\mathrm{S}'\}\frac{\sin u\xi}{u\Sigma'}\,du, \tag{5·2}$$

$$\frac{\cos n\theta_r}{\rho_r^n}$$
$$-\frac{1}{(n-1)!}\int_0^\infty \{e^{u\alpha_r} + (-)^n e^{-u\alpha_r}\}\{u\mathrm{C}' - u(\eta_r+\alpha_r)\,\mathrm{S}' + e^{-u}s\mathrm{C}'\}\frac{u^{n-1}\cos u\xi}{\Sigma}\,du$$
$$-\frac{1}{(n-1)!}\int_0^\infty \{e^{u\alpha_r} + (-)^{n+1} e^{-u\alpha_r}\}\{u\mathrm{S}' - u(\eta_r+\alpha_r)\,\mathrm{C}' + e^{-u}c\mathrm{S}'\}\frac{u^{n-1}\cos u\xi}{\Sigma'}\,du,$$
$$(n \geqslant 1) \tag{5·3}$$

$$\frac{\sin n\theta_r}{\rho_r^n}$$
$$-\frac{1}{(n-1)!}\int_0^\infty \{e^{u\alpha_r} + (-)^{n-1} e^{-u\alpha_r}\}\{u\mathrm{C}' - u(\eta_r+\alpha_r)\,\mathrm{S}' + e^{-u}s\mathrm{C}'\}\frac{u^{n-1}\sin u\xi}{\Sigma}\,du$$
$$-\frac{1}{(n-1)!}\int_0^\infty \{e^{u\alpha_r} + (-)^n e^{-u\alpha_r}\}\{u\mathrm{S}' - u(\eta_r+\alpha_r)\,\mathrm{C}' + e^{-u}c\mathrm{S}'\}\frac{u^{n-1}\sin u\xi}{\Sigma'}\,du,$$
$$(n \geqslant 1) \tag{5·4}$$

$$\begin{aligned}\frac{\cos n\theta_r+\cos(n+2)\theta_r}{\rho_r^n}&\\ +\frac{2}{n!}\int_0^\infty&\{(1-\alpha_r)e^{u\alpha_r}+(-)^n(1+\alpha_r)e^{-u\alpha_r}\}\{u(\eta_r+\alpha_r)\mathrm{S}'-u\mathrm{C}'-e^{-u}s\mathrm{C}'\}\frac{u^n\cos u\xi}{\Sigma}du\\ +\frac{2}{n!}\int_0^\infty&\{e^{u\alpha_r}+(-)^n e^{-u\alpha_r}\}\{s\mathrm{C}'-(\eta_r+\alpha_r)c\mathrm{S}'\}\frac{e^{-u}u^n\cos u\xi}{\Sigma}du\\ +\frac{2}{n!}\int_0^\infty&\{(1-\alpha_r)e^{u\alpha_r}+(-)^{n+1}(1+\alpha_r)e^{-u\alpha_r}\}\{u(\eta_r+\alpha_r)\mathrm{C}'-u\mathrm{S}'-e^{-u}c\mathrm{S}'\}\frac{u^n\cos u\xi}{\Sigma'}du\\ +\frac{2}{n!}\int_0^\infty&\{e^{u\alpha_r}+(-)^{n+1}e^{-u\alpha_r}\}\{c\mathrm{S}'-(\eta_r+\alpha_r)s\mathrm{C}'\}\frac{e^{-u}u^n\cos u\xi}{\Sigma'}du,\quad(n\geqslant 0)\end{aligned}\tag{5·5}$$

$$\begin{aligned}\frac{\sin n\theta_r+\sin(n+2)\theta_r}{\rho_r^n}&\\ +\frac{2}{n!}\int_0^\infty&\{(1-\alpha_r)e^{u\alpha_r}+(-)^{n+1}(1+\alpha_r)e^{-u\alpha_r}\}\{u(\eta_r+\alpha_r)\mathrm{S}'-u\mathrm{C}'-e^{-u}s\mathrm{C}'\}\frac{u^n\sin u\xi}{\Sigma}du\\ +\frac{2}{n!}\int_0^\infty&\{e^{u\alpha_r}+(-)^{n+1}e^{-u\alpha_r}\}\{s\mathrm{C}'-(\eta_r+\alpha_r)c\mathrm{S}'\}\frac{e^{-u}u^n\sin u\xi}{\Sigma}du\\ +\frac{2}{n!}\int_0^\infty&\{(1-\alpha_r)e^{u\alpha_r}+(-)^n(1+\alpha_r)e^{-u\alpha_r}\}\{u(\eta_r+\alpha_r)\mathrm{C}'-u\mathrm{S}'-e^{-u}c\mathrm{S}'\}\frac{u^n\sin u\xi}{\Sigma'}du\\ +\frac{2}{n!}\int_0^\infty&\{e^{u\alpha_r}+(-)^n e^{-u\alpha_r}\}\{c\mathrm{S}'-(\eta_r+\alpha_r)s\mathrm{C}'\}\frac{e^{-u}u^n\sin u\xi}{\Sigma'}du,\quad(n\geqslant 0)\end{aligned}\tag{5·6}$$

where

$$\left.\begin{aligned}s=\sinh u,\ c=\cosh u,\ \mathrm{S}'=\sinh u(\eta_r+\alpha_r),\ \mathrm{C}'=\cosh u(\eta_r+\alpha_r),\\ \Sigma=\sinh 2u+2u,\quad \Sigma'=\sinh 2u-2u.\end{aligned}\right\}\tag{5·7}$$

These functions must now be combined to make the normal and shear stresses at the edges of each hole zero. In order that this may be done each stress function must be expanded in terms of the polar co-ordinates ρ_r, θ_r at the centre O_r of a typical hole. These expansions may be obtained by straightforward algebra but the results are not given here as they are very complicated and are not likely to be used in the general case.

Stress functions for a plate bounded by one straight edge may be deduced as a limiting case from those obtained for a strip, or they may be derived independently by a method which is similar to that used for the strip.

THREE EQUAL HOLES IN A ROW

6. It was pointed out in the introduction that stress systems in a plate containing three holes in a row have special interest as experimental results exist for such a plate when it is acted on by a uniform tension. Theoretical results for this problem will now be deduced from §§ 2–4.

Let the three circles whose centres are at O_1, O_2, O_3, and the three non-dimensional co-ordinate systems

$$\zeta_1 = \xi_1 + i\eta_1 = i\rho_1 e^{-i\theta_1}, \quad \zeta_2 = \xi_2 + i\eta_2 = i\rho_2 e^{-i\theta_2}, \quad \zeta_3 = \xi_3 + i\eta_3 = i\rho_3 e^{-i\theta_3}, \tag{6·1}$$

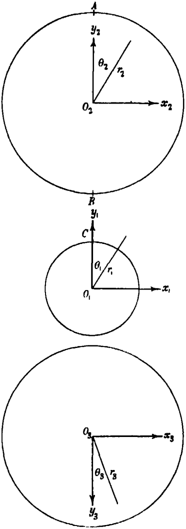


Figure 1

be as shown in figure 1, where $O_1O_2 = O_1O_3 = h$. For convenience the η_3 axis is taken in the opposite sense to the η_1 and η_2 axes and this must be remembered in making deductions from general formulae. The radii of the two outer circles will be taken to be equal and only stress systems symmetrical about both axes at O_1 will be considered so that ${}^3A_n = {}^2A_n$, ${}^3C_n = {}^2C_n$, $(n \geqslant 0)$, and the general stress function (2·5) reduces to

$$\begin{aligned}\chi = \chi_0 + {}^1A_0 \log \rho_1 + \sum_{n=1}^{\infty} {}^1A_{2n} \rho_1^{-2n} \cos 2n\theta_1 + \sum_{n=0}^{\infty} {}^1C_{2n} \rho_1^{-2n} \cos (2n+2)\,\theta_1 \\ + {}^2A_0 (\log \rho_2 + \log \rho_3) + \sum_{n=1}^{\infty} {}^2A_n (\rho_2^{-n} \cos n\theta_2 + \rho_3^{-n} \cos n\theta_3) \\ + \sum_{n=0}^{\infty} {}^2C_n \{\rho_2{}^{n} \cos (n+2)\,\theta_2 + \rho_3^{-n} (n+2)\,\theta_3\}. \qquad (6{\cdot}2)\end{aligned}$$

EXPANSIONS ABOUT THE CENTRES OF THE CIRCLES

7. As the stress system will be invariant under a transformation from the axes at O_3 to the axes at O_2 it is only necessary to consider the boundary conditions at the edges of the centre circle and at one of the outside circles, say O_2; the boundary conditions at the edge of the remaining circle will then be automatically satisfied The necessary expansions of the stress functions are given by equations (3·1)–(3·3), and, if constant terms are neglected, they are

$$\left.\begin{aligned}\log \rho_2 + \log \rho_3 &= \sum_{k=1}^{\infty} ({}^0a_{2k}^{2,1} + {}^0a_{2k}^{3,1})\, \rho_1^{2k} \cos 2k\theta_1, \\ \rho_2^{-n} \cos n\theta_2 + \rho_3^{-n} \cos n\theta_3 &= \sum_{k=1}^{\infty} ({}^na_{2k}^{2,1} + {}^na_{2k}^{3,1})\, \rho_1^{2k} \cos 2k\theta_1, \quad (n \geqslant 1) \\ \rho_2^{-n} \cos (n+2)\,\theta_2 + \rho_3^{-n} \cos (n+2)\,\theta_3 &= \sum_{k=0}^{\infty} \{{}^ne_{2k}^{2,1} + {}^ne_{2k}^{3,1} \\ &\quad + ({}^ng_{2k}^{2,1} + {}^ng_{2k}^{3,1})\, \rho_1^2\}\, \rho_1^{2k} \cos 2k\theta_1, \quad (n \geqslant 0)\end{aligned}\right\} \qquad (7{\cdot}1)$$

$$\left.\begin{aligned}\log \rho_1 &= \sum_{k=1}^{\infty} {}^0a_k^{1,2} \rho_2^k \cos k\theta_2, \\ \rho_1^{-2n} \cos 2n\theta_1 &= \sum_{k=1}^{\infty} {}^{2n}a_k^{1,2} \rho_2^k \cos k\theta_2, \quad (n \geqslant 1) \\ \rho_1^{-2n} \cos (2n+2)\,\theta_1 &= \sum_{k=0}^{\infty} \{{}^{2n}e_k^{1,2} + {}^{2n}g_k^{1,2} \rho_2^2\}\, \rho_2^k \cos k\theta_2, \quad (n \geqslant 0)\end{aligned}\right\} \qquad (7{\cdot}2)$$

$$\left.\begin{aligned}\log \rho_3 &= \sum_{k=1}^{\infty} {}^0a_k^{3,2} \rho_2^k \cos k\theta_2, \\ \rho_3^{-n} \cos n\theta_3 &= \sum_{k=1}^{\infty} {}^na_k^{3,2} \rho_2^k \cos k\theta_2, \quad (n \geqslant 1) \\ \rho_3^{-n} \cos (n+2)\,\theta_3 &= \sum_{k=0}^{\infty} ({}^ne_k^{3,2} + {}^ng_k^{3,2} \rho_2^2)\, \rho_2^k \cos k\theta_2. \quad (n \geqslant 0)\end{aligned}\right\} \qquad (7{\cdot}3)$$

The coefficients in these expansions are obtained from equations (3·4)–(3·8), and are

$$\left.\begin{aligned}
{}^{0}a_{2k}^{2,1}+{}^{0}a_{2k}^{3,1} &= 2\,{}^{0}a_{2k}^{2,1} = -\frac{1}{k}, \quad (k\geqslant 1)\\
{}^{n}a_{2k}^{2,1}+{}^{n}a_{2k}^{3,1} &= 2\,{}^{n}a_{2k}^{2,1} = 2(-)^{n}\binom{n+2k-1}{2k}, \quad (n\geqslant 1,\ k\geqslant 1)\\
{}^{n}e_{2k}^{2,1}+{}^{n}e_{2k}^{3,1} &= 2\,{}^{n}e_{2k}^{2,1} = 2(-)^{n}\binom{n+2k}{2k}, \quad (n\geqslant 0,\ k\geqslant 1)\\
{}^{n}g_{2k}^{2,1}+{}^{n}g_{2k}^{3,1} &= 2\,{}^{n}g_{2k}^{2,1} = 2(-)^{n+1}\binom{n+2k+1}{2k+1}, \quad (n\geqslant 0,\ k\geqslant 0)
\end{aligned}\right\}\tag{7·4}$$

$$\left.\begin{aligned}
{}^{0}a_{k}^{1,2} &= \frac{(-)^{k-1}}{k}, \quad (k\geqslant 1)\\
{}^{2n}a_{k}^{1,2} &= (-)^{k}\binom{k+2n-1}{k}, \quad (n\geqslant 1,\ k\geqslant 1)\\
{}^{2n}e_{k}^{1,2} &= (-)^{k}\binom{2n+k}{k}, \quad (n\geqslant 0,\ k\geqslant 2)\\
{}^{2n}g_{k}^{1,2} &= (-)^{k+1}\binom{2n+k+1}{k+1}, \quad (n\geqslant 0,\ k\geqslant 0)
\end{aligned}\right\}\tag{7·5}$$

$$\left.\begin{aligned}
{}^{0}a_{k}^{3,2} &= \frac{(-)^{k-1}}{k2^{k}}, \quad (k\geqslant 1)\\
{}^{n}a_{k}^{3,2} &= \frac{(-)^{k+n}}{2^{k+n}}\binom{k+n-1}{k}, \quad (n\geqslant 1,\ k\geqslant 1)\\
{}^{n}e_{k}^{3,2} &= \frac{(-)^{k+n}}{2^{k+n}}\binom{k+n}{k}, \quad (n\geqslant 0,\ k\geqslant 2)\\
{}^{n}g_{k}^{3,2} &= \frac{(-)^{k+n+1}}{2^{k+n+2}}\binom{k+n+1}{k+1}. \quad (n\geqslant 0,\ k\geqslant 0)
\end{aligned}\right\}\tag{7·6}$$

The coefficient of $\rho_2\cos\theta_2$ is omitted as this term contributes nothing to the stresses.

EVALUATION OF THE CONSTANTS

8. Equation (4·21) gives

$$\left.\begin{aligned}
{}^{1}A_{2n} &= \sum_{p=0}^{\infty}{}^{1}A_{2n}^{(p)}, \quad {}^{1}C_{2n} = \sum_{p=0}^{\infty}{}^{1}C_{2n}^{(p)},\\
{}^{2}A_{n} &= \sum_{p=0}^{\infty}{}^{2}A_{n}^{(p)}, \quad {}^{2}C_{n} = \sum_{p=0}^{\infty}{}^{2}C_{n}^{(p)},
\end{aligned}\right\}\tag{8·1}$$

where, from equations (4·22)–(4·24),

$$\left.\begin{aligned}
{}^1A_0^{(0)} &= -{}^1S_0 h^2\lambda_1^2,\\
{}^1A_{2n}^{(0)} &= \frac{h^2\{(n+1)\,{}^1T_{2n} - n\,{}^1S_{2n}\}\,\lambda_1^{2n+2}}{2n(2n+1)}, \quad (n \geqslant 1)\\
{}^1C'^{(0)}_{2n} &= \frac{h^2\{{}^1S_{2n+2} - {}^1T_{2n+2}\}\,\lambda_1^{2n+2}}{2(2n+1)}, \quad (n \geqslant 0)
\end{aligned}\right\} \qquad (8\cdot 2)$$

$$\left.\begin{aligned}
{}^2A_0^{(0)} &= -{}^2S_0 h^2\lambda_2^2,\\
{}^2A_n^{(0)} &= \frac{h^2\{(n+2)\,{}^2T_n - n\,{}^2S_n\}\,\lambda_2^{n+2}}{2n(n+1)}, \quad (n \geqslant 1)\\
{}^2C'^{(0)}_n &= \frac{h^2({}^2S_{n+2} - {}^2T_{n+2})\,\lambda_2^{n+2}}{2(n+1)}; \quad (n \geqslant 0)
\end{aligned}\right\} \qquad (8\cdot 3)$$

while, for $p \geqslant 0$, $n \geqslant 0$,

$$\left.\begin{aligned}
{}^1A_{2n}^{(p+1)} &= \sum_{k=0}^{\infty}\{{}^2A_k^{(p)}({}^k\alpha_{2n}^{2,1} + {}^k\alpha_{2n}^{3,1}) + {}^2C'^{(p)}_k({}^k\gamma_{2n}^{2,1} + {}^k\gamma_{2n}^{3,1})\},\\
{}^1C'^{(p+1)}_{2n} &= \sum_{k=0}^{\infty}\{{}^2A_k^{(p)}({}^k\kappa_{2n+2}^{2,1} + {}^k\kappa_{2n+2}^{3,1}) + {}^2C'^{(p)}_k({}^k\nu_{2n+2}^{2,1} + {}^k\nu_{2n+2}^{3,1})\},\\
{}^2A_n^{(p+1)} &= \sum_{k=0}^{\infty}\{{}^1A_{2k}^{(p)}\,{}^{2k}\alpha_n^{1,2} + {}^1C'^{(p)}_{2k}\,{}^{2k}\gamma_n^{1,2} + {}^2A_k^{(p)}\,{}^k\alpha_n^{3,2} + {}^2C'^{(p)}_k\,{}^k\gamma_n^{3,2}\},\\
{}^2C'^{(p+1)}_n &= \sum_{k=0}^{\infty}\{{}^1A_{2k}^{(p)}\,{}^{2k}\kappa_{n+2}^{1,2} + {}^1C'^{(p)}_{2k}\,{}^{2k}\nu_{n+2}^{1,2} + {}^2A_k^{(p)}\,{}^k\kappa_{n+2}^{3,2} + {}^2C'^{(p)}_k\,{}^k\nu_{n+2}^{3,2}\}.
\end{aligned}\right\} \qquad (8\cdot 4)$$

The coefficients in (8·4) are given by the equations (4·16)–(4·19) and (7·4)–(7·6), and are

$$\left.\begin{aligned}
{}^k\alpha_0^{2,1} + {}^k\alpha_0^{3,1} &= 0, \quad (k \geqslant 0)\\
{}^0\alpha_{2n}^{2,1} + {}^0\alpha_{2n}^{3,1} &= -\frac{2n-1}{n}\lambda_1^{4n}, \quad (n \geqslant 1)\\
{}^k\alpha_{2n}^{2,1} + {}^k\alpha_{2n}^{3,1} &= 2(2n-1)(-)^k\binom{k+2n-1}{2n}\lambda_1^{4n}, \quad (n \geqslant 1,\ k \geqslant 1)
\end{aligned}\right\} \qquad (8\cdot 5)$$

$$\left.\begin{aligned}
{}^k\gamma_0^{2,1} + {}^k\gamma_0^{3,1} &= 4(k+1)(-)^k\lambda_1^2, \quad (k \geqslant 0)\\
{}^k\gamma_{2n}^{2,1} + {}^k\gamma_{2n}^{3,1} &= 2(-)^k\lambda_1^{4n}\left\{(2n-1)\binom{k+2n}{2n} - 2n\binom{k+2n+1}{2n+1}\lambda_1^2\right\}, \quad (k \geqslant 0, n \geqslant 1)
\end{aligned}\right\} \qquad (8\cdot 6)$$

$$\left.\begin{aligned}
{}^0\kappa_{2n}^{2,1} + {}^0\kappa_{2n}^{3,1} &= 2\lambda_1^{4n-2}, \quad (n \geqslant 1)\\
{}^k\kappa_{2n}^{2,1} + {}^k\kappa_{2n}^{3,1} &= 4n(-)^{k+1}\binom{k+2n-1}{2n}\lambda_1^{4n-2}, \quad (n \geqslant 1,\ k \geqslant 1)
\end{aligned}\right\} \qquad (8\cdot 7)$$

$$ {}^{k}\nu^{2,1}_{2n} + {}^{k}\nu^{3,1}_{2n} = 2(-)^{k+1}\lambda_1^{4n-2}\left\{2n\binom{k+2n}{2n} - (2n+1)\binom{k+2n+1}{2n+1}\lambda_1^2\right\}, \qquad (k \geqslant 0,\ n \geqslant 1) \quad (8\cdot 8) $$

$$ \left.\begin{aligned} {}^{2k}\alpha^{1,2}_0 &= 0, \quad (k \geqslant 0) \\ {}^{0}\alpha^{1,2}_n &= (-)^{n-1}\frac{(n-1)}{n}\lambda_2^{2n}, \quad (n \geqslant 1) \\ {}^{2k}\alpha^{1,2}_n &= (-)^{n}(n-1)\binom{2k+n-1}{n}\lambda_2^{2n}, \quad (n \geqslant 1,\ k \geqslant 1) \end{aligned}\right\} \quad (8\cdot 9) $$

$$ \left.\begin{aligned} {}^{2k}\gamma^{1,2}_0 &= 2(2k+1)\lambda_2^2, \quad (k \geqslant 0) \\ {}^{2k}\gamma^{1,2}_n &= (-)^{n}\lambda_2^{2n}\left\{(n-1)\binom{2k+n}{n} - n\binom{2k+n+1}{n+1}\lambda_2^2\right\}, \quad (n \geqslant 1,\ k \geqslant 0) \end{aligned}\right\} \quad (8\cdot 10) $$

$$ \left.\begin{aligned} {}^{k}\alpha^{3,2}_0 &= 0, \quad (k \geqslant 0) \\ {}^{0}\alpha^{3,2}_n &= (-)^{n-1}\frac{(n-1)}{n}\frac{\lambda_2^{2n}}{2^n}, \quad (n \geqslant 1) \\ {}^{k}\alpha^{3,2}_n &= (-)^{n+k}(n-1)\binom{k+n-1}{n}\frac{\lambda_2^{2n}}{2^{k+n}}, \quad (n \geqslant 1,\ k \geqslant 1) \end{aligned}\right\} \quad (8\cdot 11) $$

$$ \left.\begin{aligned} {}^{k}\gamma^{3,2}_0 &= 2(k+1)(-)^{k}\frac{\lambda_2^2}{2^{k+2}}, \quad (k \geqslant 0) \\ {}^{k}\gamma^{3,2}_n &= (-)^{k+n}\frac{\lambda_2^{2n}}{2^{k+n}}\left\{(n-1)\binom{k+n}{n} - n\binom{k+n+1}{n+1}\frac{\lambda_2^2}{2^2}\right\}, \quad (n \geqslant 1,\ k \geqslant 0) \end{aligned}\right\} \quad (8\cdot 12) $$

$$ \left.\begin{aligned} {}^{0}\kappa^{1,2}_n &= (-)^{n}\lambda_2^{2n-2}, \quad (n \geqslant 2) \\ {}^{2k}\kappa^{1,2}_n &= (-)^{n-1}n\binom{2k+n-1}{n}\lambda_2^{2n-2}, \quad (n \geqslant 2,\ k \geqslant 1) \end{aligned}\right\} \quad (8\cdot 13) $$

$$ {}^{2k}\nu^{1,2}_n = (-)^{n-1}\lambda_2^{2n-2}\left\{n\binom{2k+n}{n} - (n+1)\binom{2k+n+1}{n+1}\lambda_2^2\right\}, \quad (n \geqslant 2,\ k \geqslant 0) \quad (8\cdot 14) $$

$$ \left.\begin{aligned} {}^{0}\kappa^{3,2}_n &= (-)^{n}\frac{\lambda_2^{2n-2}}{2^n}, \quad (n \geqslant 2) \\ {}^{k}\kappa^{3,2}_n &= (-)^{k+n+1}n\binom{k+n-1}{n}\frac{\lambda_2^{2n-2}}{2^{k+n}}, \quad (n \geqslant 2,\ k \geqslant 1) \end{aligned}\right\} \quad (8\cdot 15) $$

$$ {}^{k}\nu^{3,2}_n = (-)^{k+n+1}\frac{\lambda_2^{2n-2}}{2^{k+n}}\left\{n\binom{k+n}{n} - (n+1)\binom{k+n+1}{n+1}\frac{\lambda_2^2}{2^2}\right\}. \quad (n \geqslant 2,\ k \geqslant 0) \quad (8\cdot 16) $$

THE CONVERGENCE OF THE CONSTANTS

9. To establish the validity of this formal solution, it is necessary to prove that the series in (8·1) and (8·4) are convergent. This proof has been deposited in the archives of the Royal Society and is not printed here. The series are shown to be convergent when both λ_1 and λ_2 are less than 0·255.

TENSION PERPENDICULAR TO THE LINE OF CENTRES

10. If there were no holes a uniform tension T perpendicular to the line of centres would be given by

$$\chi_0 = \tfrac{1}{4}h^2T\rho_r^2(1+\cos 2\theta_r) \tag{10·1}$$

with stress components

$$\left.\begin{aligned} \widehat{rr}_0 &= \tfrac{1}{2}T(1-\cos 2\theta_r),\\ \widehat{r\theta}_0 &= \tfrac{1}{2}T\sin 2\theta_r,\\ \widehat{\theta\theta}_0 &= \tfrac{1}{2}T(1+\cos 2\theta_r), \end{aligned}\right\} \tag{10·2}$$

where $r = 1$ or 2 according as the origin is at O_1 or O_2.

Hence $${}^1S_0 = {}^2S_0 = {}^1T_2 = {}^2T_2 = \tfrac{1}{2}T, \quad {}^1S_2 = {}^2S_2 = -\tfrac{1}{2}T, \tag{10·3}$$

all the remaining coefficients in (4·1) being zero. Thus, from (8·2) and (8·3),

$$\left.\begin{aligned} {}^1A_0^{(0)} &= -\tfrac{1}{2}h^2\lambda_1^2T, \quad {}^1A_2^{(0)} = \tfrac{1}{4}h^2\lambda_1^4T, \quad {}^1C_0'^{(0)} = -\tfrac{1}{2}h^2\lambda_1^2T,\\ {}^2A_0^{(0)} &= -\tfrac{1}{2}h^2\lambda_2^2T, \quad {}^2A_2^{(0)} = \tfrac{1}{4}h^2\lambda_2^4T, \quad {}^2C_0'^{(0)} = -\tfrac{1}{2}h^2\lambda_2^2T. \end{aligned}\right\} \tag{10·4}$$

Calculations have been carried out for the case $\lambda_1 = \lambda_2 = 0·15$. As the tabulation of the constants which are needed in the work would take up considerable space the detailed results are omitted and only the final values of various stresses which are of interest are given here. The constants (8·5)–(8·16) have to be evaluated first and then equations (8·1)–(8·4) give the values of ${}^1A_{2n}$, ${}^1C_{2n}$, ${}^2A_n$, ${}^2C_n$. Any required stresses may then be found from the stress function (6·2).

The values of $\widehat{\theta\theta}/T$ at the edges of the centre and outside holes are shown in tables 1 and 2 respectively and the values of $\widehat{\theta\theta}/T$ and $\widehat{rr}/T$ (or $\widehat{x_1x_1}/T$, $\widehat{y_1y_1}/T$) on the line of centres between two holes are given in table 3. For comparison the values of $\widehat{\theta\theta}/T$ at the edge of one hole in an infinite plate under tension are given in table 4.

Experimental results have been obtained by Mr P. L. Capper for an arrangement of three holes of diameter d, arranged in a line across a tension member of width $9d$ with a spacing of $2d$ between the holes and of d between the outer holes and the edges of the plate, where $9d$ was taken to be 4·125 in. and the total tension on the strip per unit thickness was 3020 lb./in. This problem corresponds to the case for which $\lambda_1 = \lambda_2 = 1/6$, which are approximately the values used in the above calculations, so that some comparison of the experimental and theoretical results is possible although complete quantitative agreement cannot be expected owing to the considerable influence which the straight boundaries have on the numerical values of the stresses.

TABLE 1

θ_1°	$\widehat{\theta\theta}/T$	θ_1°	$\widehat{\theta\theta}/T$	θ_1°	$\widehat{\theta\theta}/T$
0	3·009	40	1·439	70	−0·399
10	2·896	45	1·103	80	−0·743
20	2·570	50	0·766	90	−0·863
30	2·065	60	0·126		

TABLE 2

θ_2°	$\widehat{\theta\theta}/T$	θ_2°	$\widehat{\theta\theta}/T$	θ_2°	$\widehat{\theta\theta}/T$
0	3·019	70	−0·473	130	0·752
10	2·899	80	−0·807	140	1·428
20	2·553	90	−0·915	150	2·054
30	2·024	100	−0·782	160	2·558
40	1·377	110	−0·426	170	2·884
50	0·690	120	0·108	180	2·997
60	0·047				

TABLE 3

η_1	$\widehat{\theta\theta}/T$	$\widehat{rr}/T$	η_1	$\widehat{\theta\theta}/T$	$\widehat{rr}/T$	η_1	$\widehat{\theta\theta}/T$	$\widehat{rr}/T$
0·15	3·009	0	0·4	1·150	0·265	0·7	1·257	0·333
0·2	1·793	0·388	0·5	1·126	0·243	0·8	1·784	0·385
0·3	1·262	0·336	0·6	1·148	0·264	0·85	2·997	0

TABLE 4

θ°	$\widehat{\theta\theta}/T$	θ°	$\widehat{\theta\theta}/T$	θ°	$\widehat{\theta\theta}/T$
0	3	40	1·35	70	−0·53
10	2·88	50	0·65	80	−0·88
20	2·53	60	0	90	−1
30	2				

Capper found that the stresses at the points A, B, C were $3{\cdot}21T$, $3{\cdot}07T$ and $3{\cdot}17T$ respectively which compare with $3{\cdot}019T$, $2{\cdot}997T$ and $3{\cdot}009T$

respectively, obtained from the present calculations. Thus the relative order of magnitudes of the three stresses agree and the numerical differences are probably mostly due to the influence of the boundaries. The stress $\widehat{\theta\theta}$ in the line BC has a minimum of 1·13T compared with 1·3T in the experiments. Further, the cross stress $\widehat{rr}$ in BC has two maxima, having values 0·39T at points distant 0·23h from O_1 and O_2 respectively, with a central minimum of 0·24T. The corresponding experimental results give two maxima, the one at distance 0·3h from O_2 having a value 0·34T and the one at distance 0·3h from O_1 being 0·29T, with a central minimum of 0·25T.

Howland (1935) has worked out the stress in an infinite plate under tension containing an infinite row of equal holes whose line of centres is perpendicular to the direction of the tension and it is interesting to notice that he found a maximum stress of 3·17T at the edge of a hole when the space ratio was 1/6. This compares with 3·17T in Capper's experiments and 3·009T in the present work. The maximum stress on the edge of the centre hole is altered very little by the presence of the two neighbouring holes (when $\lambda_1 = \lambda_2 = 1/6$) but it is appreciably altered by the presence of an infinite number of holes or by the presence of the two straight boundaries.

Capper found that the maximum compression at the edge of the middle hole was 0·38T, whereas 0·86T is obtained in the present work and 0·75T was found by Howland for the maximum compression at the edge of any hole of the infinite row.

The writer wishes to thank the Durham Colleges Research Committee for the use of a calculating machine.

REFERENCES

Coker, E. G. and Filon, L. N. G. 1931 *Photoelasticity* Camb. Univ. Press
Howland, R. C. J. 1929 *Proc. Roy. Soc.* A, **124**, 89
— 1930 *Phil. Trans* A, **229**, 49.
— 1935 *Proc. Roy. Soc.* A, **148**, 471.
Howland, R. C. J. and Knight, R. C. 1939 *Phil. Trans.* A, **238**, 357.
Howland, R. C. J. and Stevenson, A. C. 1933 *Phil. Trans.* A, **232**, 155.
Jeffery, G. B. 1920 *Phil. Trans.* A, **221**, 265
Knight, R. C. 1934 *Quart. J. Math.* **5**, 255.

Two-phase equilibrium in binary and ternary systems
II. The system methane-ethylene
III. The system methane-ethane-ethylene

BY M. GUTER, D. M NEWITT AND M. RUHEMANN
Dept. of Chemical Technology, Imperial College, South Kensington

(*Communicated by A. C. Egerton, Sec.R S.—Received* 6 *May* 1940)

The liquid-vapour equilibrium of the system methane-ethylene has been determined at 0, −42, −78, −88 and −104° C over a wide range of pressures and the results are shown on a pressure-composition-temperature diagram and by a series of pressure-composition curves. The liquid-vapour equilibrium of the ternary system methane-ethane-ethylene has been determined at −104, −78 and 0° C. Values for the two binary systems methane-ethane and methane-ethylene and for the ternary system methane-ethane-ethylene are shown on a composite pressure-composition diagram.

In Part I of this series (Ruhemann 1939), the results of a determination of the liquid-vapour equilibrium of the system methane-ethane at five temperatures (0, −42, −78, −88 and −104° C) over a wide range of pressures were described; it was shown, *inter alia*, that at pressures up to 15 atm. the system does not differ markedly in behaviour from that of a solution conforming to Raoult's law. The system methane-ethylene has now been investigated at the same five temperatures and over a similar range of pressures, and the two sets of data have been compared with those of a ternary system containing all three gases.

II. THE SYSTEM METHANE-ETHYLENE

The apparatus and method of determining the compositions of the co-existing phases of a binary mixture in equilibrium have already been described (1939). The methane and ethylene were either prepared in the laboratory or were purchased in cylinders. They were subsequently purified by liquefaction and fractional distillation, compressed in a water-lubricated compressor and stored in cylinders, their purity was checked at intervals by determining the ratio of carbon dioxide to steam formed on exploding samples with excess of oxygen in a constant volume apparatus. Mixtures of the required composition were prepared by condensing each component separately, at atmospheric pressure, into a thick-walled copper cylinder cooled in liquid air.

Experimental results

The equilibrium values are conveniently expressed by pressure-composition (P, x) curves which enable the compositions of the co-existing phases in equilibrium at any pressure to be read off directly and also give points on the plait-point curve of the mixture (figure 1). The critical pressures at 0, -42 and $-78°$ are 66·5, 74 and 51 atm. respectively, the P, x curves at these temperatures exhibiting regions of retrograde condensation.

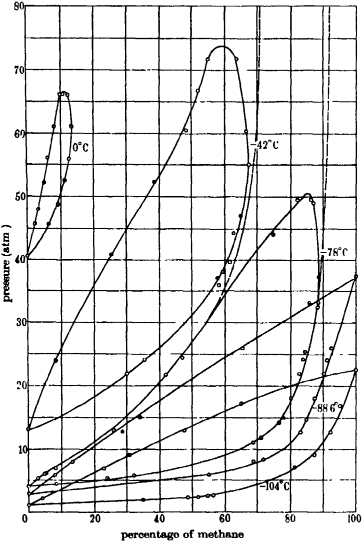


FIGURE 1

It will be recalled that in the case of the methane-ethane system at − 78° C (about 4° above the critical temperature of pure methane), the saturated vapour and boiling-point branches of the P, x curve, as determined at progressively increasing pressures, appeared to become approximately parallel at high pressures; neither curve indicated any appreciable change in composition with pressure above about 60 atm. and even up to 200 atm. there was no tendency for the curves to close. The same effect was noted at − 78° C with the methane-ethylene system (figure 1, the dotted curve).

The critical pressures of methane and ethylene are 45·8 and 50·6 atm. respectively, and although the critical pressures of mixtures of the two might be expected to be somewhat higher than that of either pure gas, such a difference as is here noted would be quite abnormal. Further work has shown that the liquid and vapour phases of both systems in the neighbourhood of the critical temperature of methane are able to exist for long periods of time in a metastable condition without measurable change in composition. True equilibrium can, however, be reached by approaching the plait-point from the direction of high pressure, the appearance of the second phase occurring in the case of methane and ethylene at about 51 atm., and in the case of methane and ethane at about 60 atm. The revised values at − 78° for the latter system are given in table 1.

TABLE 1. COMPOSITION OF LIQUID AND VAPOUR OF A METHANE-ETHANE SYSTEM IN EQUILIBRIUM AT VARIOUS PRESSURES AT − 78° C (REVISED)

	Percentages of methane by volume in	
Pressure	Liquid	Vapour
5	5·5	60·0
10	16·0	79·0
15	29·0	87·0
20	42·5	90·0
30	62·0	93·0
40	73·5	94·0
50	80·0	92·5
55	82·0	90·5
60	86·0	86·0

Temperature-composition, (T, x), *and boundary,* (P, T), *curves.* From the P, x curves shown in figure 1 values for the T, x and P, T curves of the system may be read off directly. The closed P, x curves are sections through a three-dimensional figure (figure 2) representing the variables pressure, temperature and composition of the co-existing phases, made by planes at right angles to the temperature axis. In the figure AB and CD are the vapour-pressure curves for pure methane and ethylene, respectively, BFC being the plait-point curve for the system terminating at the critical points B (for methane) and C (for ethylene). Sections through the solid by planes at right angles to the composition axis give the corresponding boundary curves (e.g. EFG is the boundary curve for a mixture containing 60 % of methane), whilst sections made by planes at right angles to the pressure axis give the T, x curves for the system (e.g. $NEPQN$ is the T, x curve at 20 atm.). It will be seen that retrograde vaporization and condensation are associated with the convexity of the solid figure with respect to two planes at right angles to one another.

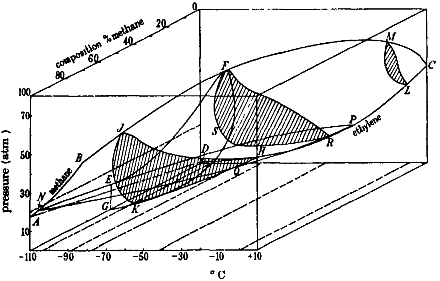


FIGURE 2

The composition of the two phases in equilibrium. The composition of the co-existing liquid and vapour phases in equilibrium are given by the points of intersection of lines parallel to the composition axis with the saturated vapour and boiling-point curves on the P, x diagram. The ratio of the concentration of either component in the vapour phase to its concentration in the liquid phase is known as the equilibrium constant, or more correctly as the equilibrium ratio; a knowledge of its value as a function of pressure or temperature is required in dealing with problems of distillation and rectification.

In the case of a mixture obeying Raoult's law if x_a, x_b and z_a, z_b are the concentrations of components A and B in the liquid and vapour phase respectively, and π_a and π_b their vapour pressures at the given temperature, then if p is the total pressure of the system

$$\frac{z_a}{x_a} = \frac{\pi_a}{p}, \quad \frac{z_b}{x_b} = \frac{\pi_b}{p}, \tag{1}$$

and at temperatures below the critical the equilibrium ratio for one component will be independent of the nature and properties of the other.

TABLE 2. THE EQUILIBRIUM RATIO OF METHANE IN THE SYSTEMS METHANE-ETHANE AND METHANE-ETHYLENE

Pressure atm	Equilibrium ratio	
	With ethylene	With ethane
	−104° C	
3	8·5	9·5
4	6·1	7·0
5	4·9	5·5
10	2·6	2·8
15	1·7	1·8
20	1·2	1·2
	−88·6° C	
8	5·4	5·5
10	4·2	4·0
15	2·7	3·0
20	1·9	1·7
25	1·5	1·4
30	1·25	1·2
35	1·05	1·05
	−77·8° C	
5	11·0	12·0
6	8·0	9·5
8	5·0	7·0
10	4·0	5·0
15	2·5	3·0
20	2·1	2·0
30	1·6	1·5
40	1·3	1·25
45	1·2	1·2
	0° C	
42	5·5	3·6
46	2·9	3·1
50	2·5	2·65
54	2·05	2·35
56	2·0	2·2
60	1·75	1·80
65	1·35	1·65

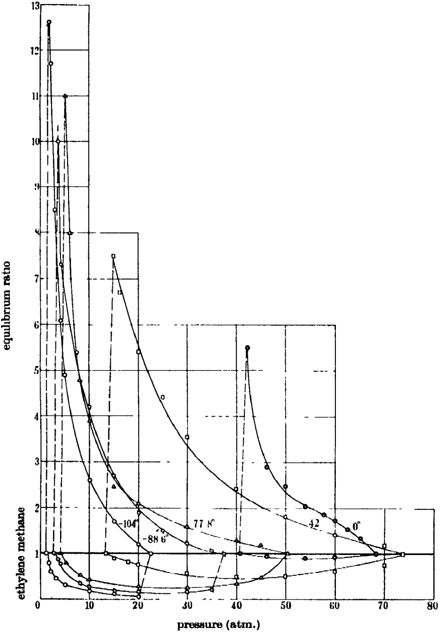

equilibrium ratio
ethylene methane
pressure (atm.)
13
12
11
10
9
8
7
6
5
4
3
2
1
0
10
20
30
40
50
60
70
80
-104°
-88 6°
77 8°
42
0°

FIGURE 3

The experimental values of z_a/x_a for methane in the systems methane-ethane and methane-ethylene are summarized in table 2. It will be seen that at the four temperatures and over the pressure range indicated in the table, the equilibrium ratios of methane in the two systems do not differ greatly. Equation (1) cannot, however, hold over the whole range of pressure since, for example, at any point on the plait-point curve the two phases become identical and $z_a/x_a = z_b/x_b = 1$.

In figure 3 the equilibrium ratios for both components of the mixture methane-ethylene are plotted as functions of the pressure. It will be noted that the ratios corresponding with points lying on or near the axes of the P, x diagram approach unity or some limiting high value, also the conditions favouring the separation of the mixture into its components are those which give high values for the methane ratio.

III. THE SYSTEM METHANE-ETHANE-ETHYLENE

In a two-phase ternary system there are three degrees of freedom, and in the experimental determination of equilibrium it is usual to fix two of them, namely, the temperature and pressure, leaving one to be satisfied. If x_a, x_b and x_c are the mole fractions of the components A, B and C and the subscripts L and V refer to the liquid and vapour phases, respectively, then

$$x_{aL} + x_{bL} + x_{cL} = x_{aV} + x_{bV} + x_{cV} = 1. \qquad (2)$$

By selecting some particular value of the mole fraction of A in the liquid phase, say x'_{aL}, then at constant temperature and pressure x'_{bL}, x'_{cL}, x'_{aV}, x'_{bV} and x'_{cV} will be fixed, and employing the conventional triangular diagram, a pair of nodes on the liquid and vapour equilibrium curves will be obtained. By varying x'_{aL} progressively a whole series of nodes can be determined and the complete equilibrium curves for both phases can be drawn. The area enclosed by the two curves on the triangular diagram gives the range of compositions within which separation into two phases can take place at the temperature and pressure to which the diagram refers.

It can easily be shown that if the ternary mixture conforms to Raoult's law the equilibrium curves will be straight lines. Thus, if p_a, p_b and p_c are the partial pressures of components A, B and C and π_a, π_b and π_c their vapour pressures in the pure state, then

$$p_a = x_{aL}\pi_a, \quad p_b = x_{bL}\pi_b \quad \text{and} \quad p_c = x_{cL}\pi_c.$$

The total pressure of the system is

$$p_a + p_b + p_c = P.$$

Let the mole fraction of component A in the liquid phase be fixed, it follows that

$$x_{bL} = \frac{P - \pi_c}{\pi_b - \pi_c} - \frac{x_{aL}(\pi_a - \pi_c)}{\pi_b - \pi_c},$$

and x_{bL} is a linear function of x_{aL}. Similar relations can be derived for x_{cL} and for the mole fractions of all the components in the vapour phase.

Experimental method

The procedure for arriving at equilibrium in the two phases and for determining the composition of the liquid and vapour is similar to that described for binary mixtures (1939), save for a modification in the method of analysis required by the presence of a third component The catharometric method of analysis determines the composition of binary gas mixtures accurately and rapidly, provided the thermal conductivities of the components differ appreciably, but it is not easily adapted for the analysis of ternary mixtures. In the present instance, however, it was found possible to remove chemically and estimate one of the gases, namely, ethylene, and to pass the remaining binary mixture through the catharometer in the ordinary way The arrangement of the apparatus for this purpose is shown in figure 4 The ternary mixture from the equilibrium vessel passes through the flow-meter F_1 and thence through three sintered glass scrubbers, two of which contain bromine dissolved in a 10 % potassium bromide solution and the third a solution of potassium bromide to remove any bromine vapour carried over with the ethylene-free gases. The residual methane-ethane mixture then passes through a second flow-meter F_2 and through the catharometer.

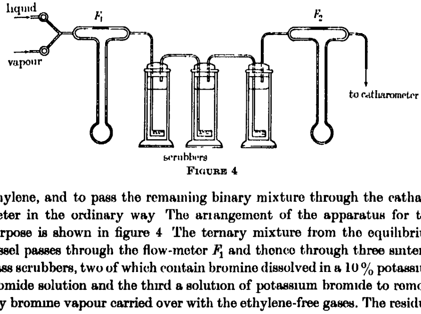


FIGURE 4

Over a limited range of velocities the flow-meter reading R is proportional to v, the amount of gas passing,

$$R = av,$$

where a is a constant the value of which depends upon the composition of the gas. It may also be assumed that for a mixture consisting of components A, B and C

$$a_m = x_a a_a + x_b a_b + x_c a_c,$$

where the subscripts m, a, b and c refer to the mixture and to its three components, respectively The values of a_a, a_b and a_c are determined experimentally The corrected readings of the first flow-meter gives the total amount of gas passing, and the difference between the corrected readings of the two flow-meters gives the amount of ethylene absorbed. In the experimental determinations of equilibrium the compositions of the mixtures were frequently checked by absorption and explosion analysis, satisfactory agreement being obtained.

Experimental results

The results at $-104°$ (at 5 and 10 atm.) and at 0° (at 50 atm.) are shown in figure 5 *a*, *b* and *c*, and those at $-78°$ (at 10, 20 and 35 atm.) in figure 6, the equilibrium points of the co-existing liquid and vapour phases being joined by connodals Comparing figure 5 *a* and *b* the effect of pressure upon the extent of the region of co-existing liquid and vapour is clearly indicated, thus, for example, at $-104°$ and 10 atm. mixtures containing upwards of 88 % of methane are entirely gaseous whilst those containing less than 34 % are liquid; at the same temperature but at 5 atm. the corresponding figures are 72 and 14 %. It is also evident that whilst the relative amounts of ethylene and ethane present in either phase vary considerably along the equilibrium curves the proportion of methane shows little change; the connodals, therefore, give the relative distribution of ethylene and ethane in the two phases for slightly fluctuating methane concentrations.

Employing equation (3) the experimental values of the concentration of a component of the mixture may be compared with the values calculated on the assumption that Raoult's law is obeyed. Thus, from the values in table 3 which refer to equilibrium at $-104°$ and 5 atm. it will be seen that the experimentally determined saturated vapour and boiling-point curves run approximately parallel to the "ideal" solution curves.

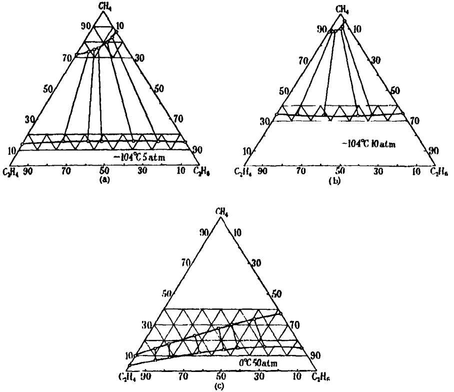


FIGURE 5

TABLE 3. COMPARISON OF THE EXPERIMENTALLY DETERMINED VALUES OF METHANE IN THE CO-EXISTING LIQUID AND VAPOUR PHASES OF A METHANE-ETHANE-ETHYLENE SYSTEM AT − 104° AND 5 ATM. PRESSURE, WITH THOSE CALCULATED FROM RAOULT'S LAW

% ethylene in liquid phase	% methane in liquid phase		% methane in vapour phase	
	Experimental	Calculated	Experimental	Calculated
62·9	15·8	19·2	74·8	86·5
50·6	15·5	19·6	76·0	88·3
42·8	16·0	19·9	76·7	89·4
28·0	16·2	20·4	80·5	92·0
13·6	16·1	20·9	83·9	93·4

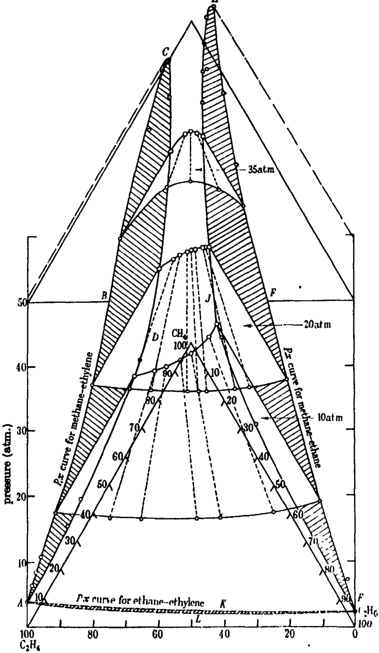


Figure 6*

* The P, x curve for ethane-ethylene (*AKEL*) shown at the base of the figure is calculated from Raoult's law.

TABLE 4. VALUES OF THE RATIOS CONCENTRATION OF ETHYLENE . CONCENTRATION OF ETHANE IN THE CO-EXISTING VAPOUR AND LIQUID PHASES OF THE SYSTEM METHANE-ETHANE-ETHYLENE

Vapour phase		Liquid phase		
% methane	Ratio ethylene/ethane $(x_b/x_e)_V$	% methane	Ratio ethylene/ethane $(x_b/x_e)_L$	Ratio $\frac{(x_b/x_e)_V}{(x_b/x_e)_L}$
		−104° C and 5 atm.		
74·8	3·94	15·8	2·95	1·34
76·0	2·33	15·5	1·50	1·56
76·7	1·74	16·0	1·05	1·66
80·5	0·76	16·2	0·50	1·51
83·9	0·34	16·1	0·19	1·77
		−104° C and 10 atm.		
88·9	6·00	34·1	5·28	1·14
89·0	2·76	33·6	1·63	1·64
90·7	1·07	33·0	0·535	2·00
91·6	0·714	33·5	0·33	2·16
		−78° C and 10 atm.		
67·9	5·17	14·9	3·65	1·42
69·1	2·92	14·4	2·05	1·43
71·3	1·577	15·1	0·956	1·64
78·7	0·625	15·5	0·673	0·93
79·5	0·322	17·7	0·237	1·36
		−78° C and 20 atm.		
84·6	4·92	37·2	3·94	1·25
86·5	3·35	37·5	2·38	1·41
87·4	1·33	37·1	1·07	1·24
88·3	0·95	37·3	0·855	1·11
88·4	0·66	37·8	0·73	0·896
88·5	0·24	38·2	0·40	0·595
88·9	0·12	38·8	0·27	0·44
		−78° C and 35 atm.		
94·3	3·07	75·2	4·63	0·663
95·0	0·723	77·9	0·99	0·73
93·9	0·22	73·2	0·218	1·01
		0° C and 50 atm.		
14·3	9·75	7·4	7·12	1·33
17·1	4·67	9·0	3·79	1·23
22·8	1·98	12·3	1·73	1·14
27·5	1·07	14·3	0·94	1·13
30·1	0·73	15·5	0·66	1·09
33·4	0·32	16·1	0·33	0·975

The relation between the two binary systems and the ternary system may be represented graphically as in figure 6 which refers to data at −78°. The equilibrium curves and their connodals at 10, 20 and 35 atm. are shown on the triangular prism the vertical axis of which refers to pressure. The P, x curves for the binary systems ($ABCD$ and $EFHJ$) will pass through the extremities of the equilibrium curves along the appropriate axis. It will be observed that they are distorted in such a way that the connodals make an angle of 60° with the horizontal.

With increase of pressure the area of the region of co-existing liquid and vapour gradually diminishes until at a pressure corresponding with the point C (50·8 atm.), which lies on the plait-point curve of the binary methane-ethylene system, the saturated vapour and boiling-point curves of the ternary mixture meet; a critical state may then be defined at the unique point at which a connodal is tangential to the curve. On progressively increasing the pressure from 50·8 atm. the region of co-existing liquid and vapour continues to diminish, and at the point H (59·9 atm.), which lies on the plait-point curve of the methane-ethane system, only one phase remains.

The conditions favourable to the separation of ethane and ethylene from the ternary mixture are seen from an inspection of the ratios of the concentration of ethylene (x_b) to the concentration of ethane (x_c) in the liquid and vapour phases at the various temperatures and pressures (table 4). In this table are also listed values of the ratio $(x_b/x_c)_V : (x_b/x_c)_L$ which give the slopes of the connodals.

The authors' thanks are due to Mr H. C. Lu for assistance with the experimental work and to the Anglo-Iranian Oil Co. for a grant towards the cost of the investigation.

REFERENCE

Ruhemann, M. 1939 *Proc. Roy. Soc.* A, **171**, 121.

BAKERIAN LECTURE

Stereochemical types and valency groups

BY N. V. SIDGWICK, F.R.S. AND H. M. POWELL

(*Delivered* 27 *June* 1940 -*Received* 27 *June* 1940)

The arrangements in space of the covalencies of a polyvalent atom, while they are subject to small variations seldom exceeding 5 or 10°, tend to conform to quite a limited number of types. It is desirable to relate this grouping to some familiar property of the atom. The property here used is the size (in G. N. Lewis's sense) of the valency group of the central atom, and the number of shared electrons which it contains, together with that of the preceding (unshared) electronic group in the atom. The experimental results show the following relations.

I. When the valency group is less than 8 we have with a covalency of 2 a linear structure (as in Cl-Hg-Cl), and with one of 3 a plane with equal angles of 120° (as in BF_3)

II. With a complete octet the arrangement can be either tetrahedral or planar. When the covalency is less than 4 it is always derived from the tetrahedron, as in the triangular OH_2 and the pyramidal NH_3. The fully shared octet is always tetrahedral when the preceding group (n in the grouping (n) 8) is 2, 8, or 18 In the transitional elements where $8 < n < 18$, it is tetrahedral if n is not much more than 8, and planar if it is not much less than 18, but the two series overlap

III. When there are 10 valency electrons, at least 2 of them (the "inert pair") unshared, the structure with a dicovalent atom (as in $M[I_3]$) is linear that of a 4-covalent atom is found in the thallous and plumbous salts to be planar, but in tellurium tetrachloride it may be a distorted tetrahedron.

IV. The peculiar 4-covalent duodecet in $M[ICl_4]$ is planar.

V. A covalency of 5 is always found to give a trigonal bipyramid.

VI. Covalency 6. Three structures are theoretically possible, a trigonal prism, a trigonal antiprism, and a regular octahedron. Experimentally the octahedron is always found, except in a few giant molecules such as those with a nickel-arsenide lattice. The regular octahedron has been found with practically every possible size of the preceding group, as well as with the "inert pair" of electrons.

VII. A covalency of 7 can have two different structures, one derived from the octahedron and the other from the trigonal prism.

VIII. Covalency 8 The only compound examined, $K_4[Mo(CN)_8]$, has a dodecahedral arrangement of the 8 CN groups.

Nearly (but not quite) all *the structures can be even more simply related* to the size of the valency group by assuming that the mean positions of the *electron pairs in this group are the same whether they are shared or not*, the structure being linear for 4 electrons, plane symmetrical for 6, either tetrahedral or plane for 8, a trigonal bipyramid for 10, and an octahedron for 12.

The experimental determination of the relative positions in space of the atoms in a molecule can now be effected by a variety of methods, of which the measurement of the diffraction of X-rays by solids and electrons by vapours, of absorption spectra and Raman spectra, and in particular cases of optical activity and electrical dipole moments, are among the most important. A knowledge of these relative positions has acquired greater importance recently for two reasons: first, because they are fundamental in determining the possibility of resonance among different structures; and secondly, from the point of view of chemical dynamics, since we now realize that the heat of activation is largely the energy required to bring the atoms into their proper positions for reaction.

On the theoretical side, the calculation of the relations between the electronic structures and the covalency directions has made great progress, mainly as the result of two methods of attack, that of localized electron pairs, and that of molecular orbitals, which are now (van Vleck and Sherman 1935) realized to be only two initial approximations leading to the same final solution. These methods enable us to relate the stereochemistry of the molecule to the number of shared and unshared electrons in the polyvalent atom, and the electronic subgroups which these occupy Recently, Kimball (1940) has given a discussion, from the group-theory point of view, of all the possible stable electronic arrangements which polyvalent atoms can have, and of the spatial relations to which they give rise.

For the chemist, who is more interested in the answer than in the method of calculation by which it is reached, these theoretical investigations are subject to two drawbacks. The first, which is inherent in the present state of molecular physics, is that when the energy differences between the various possible electronic arrangements, or between various stereochemical groupings which are compatible with the same electronic structure, are small, they cannot be determined by calculation, which thus can only inform us what arrangements are possible, and not which will actually be adopted.

The second weakness is that of the chemist, who cannot always tell whether the bonds in a given molecule are *s*, *p*, or *d* even when this is clear to the physicist, and would be glad to be able to infer the stereochemical type from some property of the molecule with which he is more familiar

It therefore seemed worth while to collect the experimental evidence as to the stereochemistry of polyvalent atoms, and to try to relate it to the simplest expression of electronic structure, the size of the valency groups and the number of shared electrons they contain.

The calculation (if it can be called so) of these sizes is simple. The total number of electrons in the combined atom is taken to be the atomic number increased by 1 for every normal covalency and by 2 for every co-ordinate covalency in which it acts as an acceptor, and further increased or diminished by the amount of the ionic charge if any. Of this total the unshared electrons are distributed as far as possible among completed electronic groups of the size $2n^2$ (2, 8, 18, 32). The shared electrons are assumed for this purpose (though this is not always true in fact) to be all in the outermost electron group; there may also be (with the transitional elements) a penultimate group of from 9 to 17 unshared electrons. When the structures of the atoms are formulated in this way, experience shows that the outermost group (the valency group proper) if it is wholly shared may contain any even number from 2 to 16; if it is partly shared (mixed) then either (*a*) it contains not more than 8 electrons in all, or (*b*) by the removal of 2 unshared electrons (known as the "inert pair") it can be reduced either to an octet or to a wholly shared group. The only certain exception to this is the anion $[ICl_4]^-$, where the valency group of the iodine contains $7+4+1 = 12$ electrons, of which only 8 are shared.

It is thus possible to express the electronic structure in this simple sense by means of the two outermost groups in the atom. for example, that of the oxygen in water as (2) 4, **4** (the shared electrons are printed in heavy type) and the divalent* palladium in py_2PdCl_2 as $-$(16) **8**.

The majority of the data we have to consider are the results of the X-ray analysis of crystal structures, and the interpretation of such results in terms of molecular linkage is sometimes a matter of uncertainty. It is not always clear whether particular links should be regarded as ionized or covalent, and it may sometimes be suspected that the atomic positions are largely distorted by internal strains of more than one kind. We have tried to avoid such difficulties by omitting these doubtful molecules, and considering only those whose structures appeared fairly certain. It should be observed that where there is reason to suspect the occurrence of resonance between ionized and covalent links (which presumably is the most frequent form of resonance), this need not prevent the atoms from having the spatial distribution required for the purely covalent structure, since it is a condition of the resonance that the positions should be nearly the same in both (or all) the resonance structures.

If these principles are clear, it remains only to consider the results.

Theory and practice agree that the arrangements in space of the covalencies of a polyvalent atom, while they are subject, from a variety of causes, to small variations seldom exceeding 5 or 10°,* tend to conform to a quite limited number of types (probably less than 12), according to the number of links, the nature of the linked atoms, and the way they are joined.

---

* The numerical value of the valency is of course taken in the sense defined by Grimm and Sommerfeld (1926) as the difference between the number of unshared electrons in the uncombined and in the combined atom.

In collecting the evidence for these structures, those molecules have been excluded in which the central atom definitely (and not merely in one resonance form) has multiple links, because they do not help us. We have little or no experimental knowledge of the effect of multiple links on the geometry of any forms other than the tetrahedral, and so from the observation of multiply linked molecules we cannot infer the original singly linked type.

The subject may be divided into two main parts. (I) structures in which the covalency of the central atom is 2, 3, or 4, and (II) those in which it is 5, 6, 7, or 8.

I. COVALENCIES OF 2, 3, AND 4

In this section we must distinguish four classes.

I. A. Valency group less than 8, i e. either the 2-covalent quartet **4**, or the 2- or 3-covalent sextet 2, **4** or **6**. To these must be added one or two molecules such as B_2H_6 and ClO_2, where the central atom falls a little short of a full octet.

I. B Valency group 8 (by far the commonest condition)· 2-covalent 4, **4**; 3-covalent 2, **6**, 4-covalent **8**. The electronic group preceding the octet in the atom may be 2, or any number from 8 to 18 inclusive; if it is more than 8 and less than 18 (transitional elements), the octet is always found to be fully shared

I. C. Valency group 10, two at least of these being unshared [(2) 4, **4**; (2) 2, **6**; (2) **8**]. This only occurs, so far as we know†, when the preceding group is 18 (the B elements of the Periodic Table), and is found especially in the heavier elements, and in the later periodic groups, the phenomenon is known as that of the "inert pair" of electrons (see Grimm and Sommerfeld 1926).

I. D. Valency group 12, of which 4 are unshared. this is only known for certain in one molecule, the univalent ion $[ICl_4]^-$.

---

* Occasionally the internal strain causes large distortion, as in cyclopropane and phosphorus vapour P_4, in which the valency angles are reduced to 60°.

† A possible exception, which, however, does not concern us here, is the tetravalent uranous ion U^{4+}.

I. A. *Less than* 8 *electrons in the valency group*

Quartet, 2-covalent, **4**: this is always linear. Examples are given by argentous, aurous, mercurous, mercuric, and thallic compounds: e.g. $K[Ag(CN)_2]$ (Hoard 1933 X)*, $(CH_3)_3P \rightarrow Au—I$ (Mann, Wells and Purdie 1937 X), Cl-Hg-Hg-Cl (Havighurst 1925 X); Br-Hg-Br (Braekken and Scholten 1934 X), $[CH_3\text{-}Tl\text{-}CH_3]X$ (Powell and Crowfoot 1932 X).

Sextet. The fully shared (3-covalent) sextet has only been examined in the trihalides and trimethide of boron (Levy and Brockway 1937 E), all of which have been shown to be plane structures with a covalency angle of 120°.

An interesting modification of the sextet, of which only a few examples are known, is the dicovalent form 2, **4**, this occurs in the covalent state of the plumbous and stannous halides such as $SnCl_2$, $PbBr_2$. We should expect from the behaviour of the 2- and 3-covalent octets that the valency angle would be much the same here as when the sextet is fully shared, i.e nearer to 120 than 180°: unpublished work by Sutton and Lister on the electron diffraction by the vapour of stannous and plumbous halides indicates that this is so.

Two other forms of "imperfect octet" are found in diborane, B_2H_6, and in chlorine dioxide, ClO_2, if we assume that the valency electrons are equally distributed among the covalencies, the boron in B_2H_6 has 48/7 and the chlorine in ClO_2 23/3 electrons. The structures, as we might expect, are much the same as though the octets were complete, diborane having the structure of ethane (Mark and Pohland 1925 X, Bauer 1937 E), and chlorine dioxide a valency angle which is nearer to 110 than 180° (Sutton and Brockway 1935 E).

I. B. *The octet*

(*a*) Covalency 2: 4, **4** Examples of this are numerous especially in compounds of oxygen and the sulphur group (O, S, Se, Te, Cl, Br), the valency angle is practically always between 90 and 110°, the mean being about 105° (see list in Pauling (1939, p. 79)) †

(*b*) Covalency 3: 2, **6**. This is invariably pyramidal, the valency angles being, as with the dicovalent octet, between 90 and 110°, and usually about 105°; this has been established for the 11 elements N, P, As, Sb, Bi, O, S, Se, Te, Br, I.

---

* The letters following the date indicate the method of measurement used thus X = X-ray crystal analysis: see below, p. 166.

† In some crystalline structures a linear arrangement of the dicovalent octet appears to be possible. Thus in the pyrophosphate ion P_2O_7 the P-O-P group was found to be linear (Levi and Peyronel 1936 X), and so too is the Si-O-Si group in Si_2O_7 (Gossner and Mussgnug 1929; Zachariasen 1930*b*), and in cristobalite (Wyckoff 1925*a*), though not in quartz. The Be-F-Be valencies in crystalline beryllium fluoride also seem to be nearly linear (Brandenberger 1932 X; Warren and Hill 1934 X). Whether this is due to rotation or strain in the crystal or to other causes is uncertain.

(*c*) Covalency 4: **8**; fully shared octet. This is the commonest of all covalent structures. It can assume either of two quite different stereochemical forms, the tetrahedron or a plane structure with the valencies at right angles. Both of these have long been recognized; the tetrahedron was proposed for carbon by van't Hoff nearly 70 years ago (1874), the planar form was first suggested for molecules of the type of $Pt(NH_3)_2Cl_2$ by Werner some 20 years later (1893), this view of Werner's has been disputed, but is now established beyond doubt It has been shown (Pauling 1931, Kimball 1940) that the type of structure adopted depends on the nature of the linking electrons, sp^3 and d^3s giving the tetrahedron, and dsp^2 and d^2p^2 the plane square, but its relation to the simpler chemical classification here adopted is uncertain. It obviously depends at least in part on the size of the preceding electronic group; when this contains 2 electrons (Be, B, C, N, O) or 8[†] (Mg, Al, Si, Ti, P, V, Nb, Ta, S, Cr, Mo, W, Cl, Mn, Re, Os) or 18[*] (Cu, Ag, Zn, Cd, Hg, Ga, In, Tl, Ge, Sn, Pb, As, Sb, Te, I, Fe^{-2}, Co^{-1}, Ni^0) the type is tetrahedral[†]: some 40 elements with the structures (2) **8**, (8) **8**, and (18) **8** have been shown to give tetrahedral molecules.

But when the atom has a transitional core, with an outer group (next to the octet) of more than 8 but less than 18 electrons, both of the types are found, the tetrahedral when the group is not much more than 8, and the planar when it is not much less than 18, and the two series overlapping in the middle As to several of these structures we unfortunately have no data, but enough have been determined to show the general relations, they are given in the following table, where *n* is the number of electrons in the (unshared) group preceding the octet:

n =	9	10	11	12	13	14	15	16	17
Tetrahedral	$\mathrm{Mn^{vi}}$ $\mathrm{Re^{vi}}$				$\mathrm{Mn^{ii}}$ $\mathrm{Fe^{iii}}$	$\mathrm{Pt^{iv}}$	$\mathrm{Co^{ii}}$		
Planar					$\mathrm{Mn^{ii}}$		$\mathrm{Co^{ii}}$	$\mathrm{Au^{iii}}$ $\mathrm{Ni^{ii}}$ $\mathrm{Pd^{ii}}$ $\mathrm{Pt^{ii}}$	$\mathrm{Cu^{ii}}$ $\mathrm{Ag^{ii}}$

---

[*] All these elements are in their group valencies unless otherwise marked.

[†] The only exception to this is found in the aurous compounds $K[Au(CN)_2A]$, where *A* stands for a molecule of the chelate diamine dipyridyl or phenanthroline; Wardlaw and his colleagues (Dothie, Llewellyn, Wardlaw, and Welch 1939) claim to have shown that these are planar, although the gold atom has the structure (18) **8**.

Of the atoms with the group (14) **8** only one has been measured, $(CH_3)_3PtCl$ (Cox and Webster 1935*b* X), which was found to be tetrahedral. The group (16) **8** includes a large number of compounds of divalent nickel, palladium, and platinum, whose structures have been examined in great detail They are always found to be planar (see especially Cox, Pinkard, Wardlaw and Webster 1935, Cox, Wardlaw and Webster 1935); and it seems probable that all the atoms with this grouping are of the planar type*.

This remarkable divergence of type may perhaps be explained as follows The square planar configuration is that of an octahedron with 2 *trans*-groups removed, and it may possibly involve a 4-covalent duodecet, 4, **8**, as we have in the planar $[ICl_4]^-$ ion If so, the electronic arrangement should be written not (*n*) **8** but (*n* − 4) 4, **8**, and since the penultimate group cannot be reduced below 8, the planar structure would only become possible when *n* is at least 12, which is roughly where it first appears.

I. C *The decet*

This section comprises the 2-, 3- and 4-covalent molecules of the "inert pair" kind. (The 5-covalent decets are discussed later.)

The dicovalent form (2) 4, **4** is linear, as is shown in the trihalide ions of the salts $M[ICl_2]$, $M[IBrCl]$, $M[I_3]$ (Wyckoff 1920, Mooney 1935*a*, *b*, 1938*a*). The tricovalent state (2) 2, **6** occurs in the iodide chlorides $Ar.ICl_2$, but their structure has not been examined. The 4-covalent form (2) **8** was examined by Cox (Cox, Shorter and Wardlaw 1937, 1938 X) in compounds of monovalent thallium $[Tl(SC(NH_2)_2)_4]NO_3$ and of divalent lead (a variety of plumbous complexes including the complex oxalate $K_2[Pb(C_2O_4)_2]$), all of which were found to be planar.

It is not certain, however, that the group always has this form. Tellurium tetrachloride ($TeCl_4$. − (18) (2) **8**) was found by Smyth, Grossman and Ginsburg (1940) to have in benzene solution at 25° the large dipole moment of 2·54 D. If this is the true moment of the $TeCl_4$ molecule, it cannot have either the tetrahedral or the plane structure Kimball and Pauling (quoted by Smyth *et al.* 1940) consider that this is so· that as the tellurium atom here has 10 valency electrons the structure should be that of a trigonal bipyramid (as in the 5-covalent molecules) with one place—probably equatorial—unoccupied; this would give a polar molecule of the shape of an unsymmetrical tetrahedron*.

---

* This, however, is not certain. The planar 4-covalent nickelous compounds have been shown to be diamagnetic, as Pauling's theory requires; but there are other 4-covalent nickelous compounds which have been found to be paramagnetic, although their stereochemical structure has not been examined.

On the other hand, it is possible that tellurium tetrachloride has the same planar structure as the thallous and plumbous compounds examined by Cox, and that the apparent moment is due either to abnormal atomic polarization or to the formation of a co-ordinate link to the solvent. Various facts support this view. Tellurium tetrachloride is by no means a simple covalent molecule; it melts at 225° and boils at 390°: its specific conductivity in the fused state is that of a salt, being 0·1145 at 236° (Biltz and Voigt 1924). Though it dissolves in benzene (about 4 % at the ordinary temperature), it is too insoluble in carbon tetrachloride, carbon bisulphide, and heptane for these solvents to be used, which strongly suggests that it is solvated in benzene. There are many parallel instances Mercuric chloride, bromide, and iodide, which in the vapour have been shown to be non-polar linear molecules (Braune and Linke 1935 D), are nearly insoluble in benzene, but readily dissolve in dioxane, in which solvent they were found to give moments of 1·29, 1·06, and 0·58 D respectively (Curran and Wenzke 1935 D); in this case compounds with dioxane Hg(hal)$_2$, $C_4H_8O_2$ have since been isolated (Crenshaw, Cope, Finkelstein and Rogan 1938). Aluminium bromide, which is non-polar in carbon bisulphide, has a large moment (4·89 D) at high dilutions in benzene (Ulich and Nespital 1931,) iodine has a moment of 1·0 in benzene, but is non-polar in hexane and cyclohexane (Müller and Sack 1930).†

I. D.

This structure 4, 8 is a duodecet with four unshared electrons, as though it had two inert pairs. It occurs in the tetrachloroiodides M[ICl_4]; the potassium and tetramethyl ammonium salts have been examined (Mooney 1938*b*), and it has been shown that the anion is planar and equiangular. This is the only known example of this type of electronic grouping, unless, as is possible, it also occurs in the planar 4-covalent compounds of the later transitional elements.

---

* L. Helmholtz and M T. Rogers (*J Amer Chem Soc* 1940, 62, 1537) have recently found that the anion of K[IO_2F_2], in which the iodine has the valency groups —(18) (2) 8, has a configuration that may be regarded as that of a trigonal bipyramid with one equatorial position unoccupied.

† Since this was written Stevenson and Schomaker have published (1940) an examination of tellurium tetrachloride by electron diffraction. Their results show the Te-Cl distance to be 2·33 A (sum of covalent radii 2 36), which requires that the Cl···Cl distance should be for a tetrahedron 3·82 A, and for a plane square structure 3·30 A The observed value of the principal Cl· ·Cl distance is 3·27 A, corresponding to a valency angle of 93 ± 3°. They consider that this is evidence of a distorted tetrahedron of the kind contemplated above; but it will be seen that the results lie very close to those required for a plane symmetrical structure.

II.

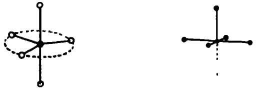
This section comprises the atoms with covalencies of 5, 6, 7, and 8.

II. A. *Covalency* 5

Several compounds of this kind have recently been measured, belonging to 4 different groups

(*a*) – (8) **10**; PF_5, PCl_5, $NbCl_5$; $NbBr_5$; $TaCl_5$; $TaBr_5$.
(*b*) – (9) **10**; $MoCl_5$.
(*c*) – (16) **10**; $Fe(CO)_5$.
(*d*) – (18) **10**, $Sb(CH_3)_3(hal)_2$.

All were found to be trigonal bipyramids (figure 1); in PCl_5 two different investigators (Rouault 1938*b*, Schomaker quoted by Pauling 1939, p 103 note) found that the polar P-Cl distances were about 0·1 A shorter than the equatorial, though the values obtained in the two investigations differed somewhat, in none of the other AB_5 compounds (types *a*, *b*, *c*) was any indication of such differences found, but the results do not always make it possible to determine this with any exactness. The last compound on the list, $Sb(CH_3)_3(hal)_2$, has the two halogen atoms in the polar positions, and at distances from the central antimony atom intermediate between those required for an electrovalent and for a covalent bond, indicating resonance between those two forms; it may be pointed out that in the ionized form $[(CH_3)_3Sb]hal_2$ the antimony has only a sextet of valency electrons, so that the natural positions of the methyl groups are the same in both structures.



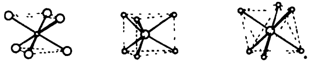
trigonal pyramid IF$_5$ structure

FIGURE 1

A fifth molecule of this covalency, IF_5, has also been examined by electron diffraction; but the difference between the atomic numbers of the atoms (9, 53) makes it impossible to determine the stereochemical type. This is peculiarly unfortunate, because there are theoretical reasons to expect that since the iodine has 12 valency electrons this molecule will be a square pyramid (figure 1).

II. B. *Covalency 6*

Next to that of 4, this is the covalency that is most often found. On theoretical grounds we should expect three stereochemical types: (1) a trigonal prism, (2) a trigonal antiprism, and (3) a regular octahedron (figure 2). In fact, it is found that all the known molecules of the AB_6 kind form regular octahedra, with the exception of a few giant molecules containing AB_6 groups, which are of the other forms (MoS_2, trigonal prism: crystals with the nickel arsenide lattice, trigonal prisms and trigonal antiprisms). The reasons for this are clear if we examine the structures more closely. Properly speaking these should be called two types, (1) the trigonal prism and (2) the trigonal antiprism: the regular octahedron is a species of (2) in which the B-B distances and the valency angles are equal. The relation between the prism and the antiprism is exactly like that between the unstable and the stable positions in ethane, and it is natural that in molecules of the AB_6 kind, where the B atoms are crowded together, they (or their links) should repel one another, which would make the antiprism form more stable. For the same reason we should expect that among the various forms of the antiprism the strain would be least in the octahedron, where the B atoms are equidistant.



trigonal antiprism trigonal prism octahedron

FIGURE 2

Hence it is natural that the octahedral form should be preferred when the AB_6 molecule is left to itself, either in the vapour or in those crystals where it is only attached to its neighbours by weak van der Waals forces; while in giant molecules the much stronger valency forces acting on its atoms may draw them over into one of the other forms.

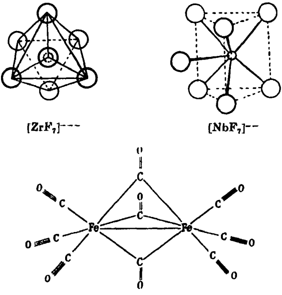
The range of the octahedral type is very wide; for atoms in the first short period a covalency of 6 (= (2) **12**) is of course impossible, but almost every other conceivable form is known, not only - (*n*) **12** with every value of *n* from 8 to 18, but also the form with the inert pair – (18) (2) **12** The elements concerned are given in the following list:

(8) **12**: Mg, Al, Si, Ti, Zr, Hf, P, V, S, Mo, W, U.
(9) **12**: [apparently the only form that is missing].
(10) **12**: V^{III}, Os^{VI}.
(11) **12**: Cr^{III}, Mo^{III}, W^{III}.

(12) **12**: $\mathrm{Ru^{iv}}$, $\mathrm{Os^{iv}}$.
(13) **12**: $\mathrm{Mn^{ii}}$, $\mathrm{Fe^{iii}}$, $\mathrm{Ru^{iii}}$, $\mathrm{Ir^{iv}}$.
(14) **12**: $\mathrm{Cr^{0}}$, $\mathrm{Mo^{0}}$, $\mathrm{W^{0}}$, $\mathrm{Fe^{ii}}$, $\mathrm{Co^{iii}}$, $\mathrm{Ru^{ii}}$, $\mathrm{Rh^{iii}}$, $\mathrm{Pd^{iv}}$, $\mathrm{Ir^{iii}}$, $\mathrm{Pt^{iv}}$.
(15) **12**: $\mathrm{Co^{ii}}$.
(16) **12**: $\mathrm{Ni^{ii}}$.
(17) **12**: $\mathrm{Cu^{ii}}$.
(18) **12**· Zn, Cd, Ga, In, Tl, Ge, Sn, Pb, As, Se, Te, I.
(18) (2) **12**: "inert pair" $\mathrm{Se^{iv}}$, $\mathrm{Te^{iv}}$.

II C *Covalency* 7

This covalency is very rare, and so far only three or perhaps four examples of it have been examined. The heptafluorozirconate $K_3[ZrF_7]$ has been shown (Hampson and Pauling 1938) to be truly 7-covalent, the structure of the anion being derived from an octahedron by adding a fluorine atom at the centre of one face, on the other hand, the niobium and tantalum compounds $K_2[NbF_7]$ and $K_2[TaF_7]$ have a different structure (Hoard 1939), derived from the trigonal prism by adding a fluorine atom at the centre of a prism face (figure 3) This difference is the more remarkable since the central atoms in the two groups of compounds are isoelectronic, having the structure – (8) **14**, though the ionic charges are of course different.



$[ZrF_7]^{---}$

$[NbF_7]^{--}$

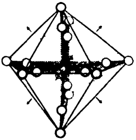
$Fe(CO)_9$

FIGURE 3

Another molecule assigned by Pauling, no doubt correctly, to the 7-covalent class is the very peculiar iron enneacarbonyl $Fe_2(CO)_9$ (Powell and Ewens 1939). The two iron atoms are joined through three carbon atoms, but their distance apart makes it almost certain that they are also linked directly, a conclusion which is supported by the diamagnetism of the compound (figure 3). Pauling (1939, p. 238) describes the iron here as trivalent, but with this formula it is tetravalent, and the valency group is – (12) 14. The stereochemical type is the same as that of the heptafluorozirconate.

II. D. *Covalency* 8

Compounds of this covalency are much more frequent than the 7-covalent, but only one of them has been examined, and that is the very stable octocyanide of tetravalent molybdenum $K_4[Mo(CN)_8]$, the anion of which was found (Hoard and Nordsieck 1939) to have neither the cubic nor the antiprismatic structure, but that of a dodecahedron (figure 4) The cube, which has generally been assumed to be likely to be the shape of an 8-covalent complex, is on theoretical grounds unlikely to occur.



$[Mo(CN)_8]^{----}$

FIGURE 4

CONCLUSION

If we compare the stereochemical type with the size of the valency group as a whole, and assume that the electron pairs occupy much the same positions whether they are shared or not, it is seen that this size is closely related to, and in most cases uniquely determines, the type of spatial arrangement adopted, and in the following way.

With a quartet of electrons, the molecule is linear (as in Cl-Hg-Cl).

With a sextet, the arrangement is planar, and the valency angles 120°, giving with a covalency of 3 the plane symmetrical molecule (as in BF_3) and where it is 2-covalent, as in $SnCl_2$, a triangular molecule.

With an octet there appear to be two different types, the tetrahedron and the plane square form. All the 2- and 3-covalent octets are derived from the tetrahedron, the 2-covalent being triangular and the 3-covalent pyramidal. The 4-covalent (fully shared) octet is tetrahedral when the electron group (*n*) preceding the valency octet contains 2, 8, or 18 electrons, and also in the earlier transitional elements when *n* is not much more than 8. When *n* is not much less than 18 the planar form appears, while for intermediate values both types occur. The planar form may really be a 4-covalent duodecet, as in the $[ICl_4]^-$ ion (below), which would become possible when *n* is not less than 12.

The decet when fully shared (5-covalent) gives the trigonal pyramid. The 4-covalent decet seems to have two forms. (1) a plane square structure in thallous and plumbous compounds (involving a change in the disposition of the 10 electrons), and perhaps (2) in tellurium tetrachloride a distorted tetrahedron, which would naturally arise from the bipyramid by the removal of one of the attached atoms. The 2-covalent decet is found in the trihalide ions such as $[I_3]^-$, which have been shown to be linear, as if derived from the trigonal bipyramid by removing all three equatorial groups.

The duodecet when fully shared (6-covalent) is octahedral. The 5-covalent form occurs in IF_5, but its structure is not known; the 4-covalent form is that of the iodine in the anion $[ICl_4]^-$, which has been shown to be square, and so to be derived from the octahedron by the removal of two trans-groups; the same structure may occur in the planar 4-covalent compounds of the later transitional elements.

The 14-group occurs in two 7-covalent forms derived respectively from the octahedron and from the trigonal prism by the insertion of a seventh valency.

Of the 16-group (8-covalent) only one example has been examined, which is found to be dodecahedral.

We wish to thank Dr Hampson and the Chemical Society, and Dr Hoard and the American Chemical Society, for permission to reproduce figures 3 and 4.

BIBLIOGRAPHY

The bibliography of this subject is very large, and only a selection of the more important papers is given here. In the first section the short titles of the papers are quoted under the elements in question, the full references being given in § 2. The elements in § 1 are arranged under their periodic groups, the symbol being followed by a Roman figure giving the valency, and an Arabic the covalency: the composition of the outer electronic groups then follows. Thus for heptavalent 4-covalent manganese (as in $KMnO_4$) we have Mn VII 4: – (8) **8**. The letters following the date of each paper show the method of observation employed: X = X-ray crystal analysis; Cr = crystallography without the use of X-rays (these two headings include papers which establish the similarity of one crystal to another whose structure has been determined in detail elsewhere); E = electron diffraction; S = spectrum, D = electrical dipole moment; O = optical activity; I = geometrical isomerism; more general papers are not marked. Tetr. = tetrahedral; oct. = octahedral.

Group I

Copper. Cu I 4: – (18) **8**: tetr. CuCl, etc.: Pauling 1939, p. 165 X. $K_3[Cu(CN)_4]$ and $[Cu(S{=}C(CH_3)NH_2)_4]Cl$: Cox, Wardlaw and Webster 1936 X. See also Mann, Purdie and Wells 1936 X; Wells 1936*a* X; Mellor, Burrows and Morris 1938 I; Ketelaar 1931 X.

Cu II 4: – (17) **8**. planar. $CuCl_2, 2H_2O$ Harker 1936 X. Cu″ diketones: Cox and Webster 1935*a* X. See also Cox, Sharrat, Wardlaw and Webster 1936 X; Cox, Wardlaw and Webster 1936 X; Mellor, Burrows and Morris 1938 I.

Cu II 6: – (17) **12**: oct. $M_2[CuCl_4, 2H_2O]$: Hendricks and Dickinson 1927 X; Chrobak 1934 X. See also Wahl (1927*b* O), with objections of Johnson and Bryant (1934) and Mandal (1939).

Silver. Ag I 2: – (18) **4**: linear. AgCN: West 1935*a* X. $K[Ag(CN)_2]$: Hoard 1933. $[Ag(NH_3)_2]_2SO_4$: Corey and Wyckoff 1934 X. Cf. Elliott and Pauling 1938 X.

Ag I 4. – (18) **8**: tetr. AgI: Bloch and Moller 1931 X; Helmholz 1935 X. $(Et_3As \rightarrow Ag{-}I)_4$: Mann, Wells and Purdie 1937 X; Wells 1936*a* X. Cf Ketelaar 1931 X; Hein and Regler 1936 O; Cox, Wardlaw and Webster 1936 X.

Ag II 4: – (17) **8**: planar. Ag″ picolinate: Cox, Wardlaw and Webster 1936 X.

Gold. Au I 2: – (18) **4**: linear. $[AuCl_2]^-$ ions in $Cs_2AuAuCl_6$: Elliott and Pauling 1938 X. $(Alkyl)_3P \rightarrow Au{-}I$: Mann, Wells and Purdie 1937 X

Au I 4: – (18) **8**: planar. $K[Au(CN)_2A]$ (A = dipyridyl and phenanthroline): Dothie, Llewellyn, Wardlaw and Welch 1939 X.

Au III 4: – (16) **8** planar. $K[AuBr_4], 2H_2O$: Cox and Webster 1936 X. $(Alk_2AuBr)_2$ and $(Alk_2AuCN)_2$. Burawoy, Gibson, Hampson and Powell 1937 X. Cf. Elliott and Pauling 1938 X; Phillips and Powell 1939 X; Gibson 1939.

Group II

Beryllium. Be II 4: (2) **8**: tetr. BeS, etc.: Goldschmidt 1926*a* X, 1927 X. $M_2[BeF_4]$: Goldschmidt 1926*b*, *c* X; Zachariasen 1926*b* X; Hultgren 1934 X. $[Be(OH_2)_4]SO_4$: Beevers and Lipson 1932 X. Dicholate Be″: Burgess and Lowry 1924 O: Mills and Gotts 1926 O. BeF_2: Brandenberger 1932 X; Warren and Hill 1934 X.

Magnesium. Mg II 4: – (18) **8**: tetr. MgTe: Zachariasen 1927 X.

Mg II 6: – (8) **12**: oct. $[Mg(OH_2)_6]SiF_6$: Hassel 1927 X; Hassel and Salvesen 1927 X.

Zinc. Zn II 4: – (18) 8: tetr. ZnS, etc.: Goldschmidt 1927 X; Zachariasen 1926*a* X. $Zn(NH_3)_2Cl_2$: MacGillavry and Bijvoet 1936 X. $K_2[Zn(CN)_4]$: Dickinson 1922*b* X. Zn benzoyl-pyrotartrate: Mills and Gotts 1926 O. Cf. Wyart 1926 X.

Zn II 6: – (18) 12: oct. $[Zn(OH_2)_6]hal_2$: Yü and Beevers 1936 X. Cf. Pfeiffer and Quehl 1931 O, 1932 O.

Cadmium. Cd II 4: – (18) 8: tetr. CdTe: Zachariasen 1926*a* X. $K_2[Cd(CN)_4]$: Dickinson 1922*b* X. $[Cd(NH_3)_4](ReO_4)_2$· Pitzer 1935 X.

Cd II 6: – (18) 12: oct. $[Cd(dipy)_3]X_2$, etc.: Pfeiffer and Quehl 1931 O, 1932 O. Cf. Neogi and Mukherjee 1934 O.

Mercury. Hg II 2: – (18) 4: linear. Mercurous: Hg_2Cl_2, Hg_2Br_2, Hg_2I_2: Havighurst 1925 X, 1926 X; Mark and Steinbach 1926 X; Huggins and Magill 1927 X. Mercuric· $HgCl_2$, $HgBr_2$, HgI_2: Braekken and Scholten 1934 X; Wehrli 1937 S; Braune and Knoke 1933 E, Gregg, Hampson, Jenkins, Jones and Sutton 1937 E; Braune and Linke 1935 D in vap. See also Hampson 1934 D.

Hg II 4: – (18) 8: tetr. HgS: Kolkmeijer, Bijvoet and Karssen 1924 X. HgTe Zachariasen 1926*a* X. $Ag_2[HgI_4]$: Ketelaar 1931 X. $K_2[Hg(CN)_4]$: Dickinson 1922*b* X.

GROUP III

Boron: always trivalent, (1) with 3-covalent sextet as in BF_3, (2) with 4-covalent "reduced octet" (strictly 48/7 electrons) as in B_2H_6; (3) with 4-covalent octet, as in $M[BF_4]$.

(1) Sextet (2) 6: flat, 120°. BF_3, BCl_3, BBr_3, $B(CH_3)_3$: Levy and Brockway 1937 E: Gregg, Hampson, Jenkins, Jones and Sutton 1937 E. BF_3· Linke and Rohrmann 1937 D. BCl_3: Ulich and Nespital 1931*b* D. Cf. Bailey, Hall and Thompson 1937 S, Zachariasen 1934 X.

(2) (2) 48/7: tetr. B_2H_6: Mark and Pohland 1925 X, Bauer 1937 E.

(3) (2) 8: tetr. Borosalicylates: Boeseken and Meulenhoff 1924 O. $M[BF_4]$: Vorländer, Hollatz and Fischer 1932 Cr; Hoard and Blair 1935 X. Cf. Keggin 1933 X, 1934 X; Klinkenberg and Ketelaar 1935 X; Zachariasen 1937 X.

Aluminium. Al III 4: – (8) 8: tetr. AlP, AlN, etc.: Pauling 1939, p. 165. Al_2Cl_6, Al_2Br_6· Palmer and Elliott 1938 E.

Al III 6: – (8) 12: oct. Complex oxalate. Wahl 1927*a* O. Catechol complex: Treadwell 1932 O. Na_3AlF_6: Náray-Szabó and Sasvári 1938 X. $(NH_4)_3AlF_6$: Pauling 1924 X, Menzer 1930 X.

Gallium. Ga III 4: – (18) 8: tetr. GaP, GaAs, GaSb: Pauling 1939, p. 165 X; Goldschmidt 1927 X. GaN: Juza and Hahn 1938 X.

Ga III 6: – (18) 12: oct. $M_3[Ga(C_2O_4)_3]$: Neogi and Dutt 1938 O.

Indium. In III 4: – (18) 8: tetr. InN: Juza and Hahn 1938 X. InSb: Pauling 1939, p. 165 X.

Thallium. Tl III 2. – (18) 4: linear. $[Tl(CH_3)_2]I$: Powell and Crowfoot 1934 X.

Tl III 4: – (18) 8. tetr. Acetylacetonate $(CH_3)_2TlA$· Cox, Shorter and Wardlaw 1938 X Cf. Sidgwick and Sutton 1930.

Tl III 6: – (18) 12: oct. $K_3[TlCl_6]$, $2H_2O$· Hoard and Goldstein 1935*c* X. $M_3[Tl_2Cl_9]$: Powell and Wells 1935*b* X: Hoard and Goldstein 1935*b* X.

Tl I 4: – (18) (2) 8: planar $[Tl(S{=}C(NH_2)_2)_4]NO_3$· Cox, Shorter and Wardlaw 1938 X. Phenanthroline complex, $[Tl(phen)_2]NO_3$: Wardlaw 1940 X unpublished.

GROUP IV

Carbon. C IV 4: (2) **8** –: always tetr.· van't Hoff 1874 O. Diamond: Bragg, W. H. and W. L. 1913 X; Bragg, W. H. 1921 X; Ehrenberg 1926 X. Tetrahalides: Finbak and Hassel 1937 X; Brockway and Wall 1934 E; Degard, Pierard and v.d. Grinten 1935 E and X-ray on vapour.

Silicon. Si IV 4: – (8) **8**: tetr. Si: Manchot 1922 X. α-quartz: Bragg, W. H. and Gibbs 1925 X; Gibbs 1926 X. SiS_2: Bussem, Fischer and Gruner 1935 X. SiF_4, $SiCl_4$: Brockway and Wall 1934 E SiI_4: Hassel and Kringstad 1931 X. $Si(CH_3)_4$: Brockway and Jenkins 1936 E. $Si(C_6H_5)_4$· Giacomello 1938 X. $SiHCl_3$. Pirenne 1938, 1939 X-ray on vapour; Hemptinne and Wouters 1936 S Opt. active silicanes, etc.: Kipping 1907, 1908 O.

Si IV 6: – (8) **12**. oct. $M_2[SiF_6]$: Ketelaar 1935 X.

Germanium. Ge IV 4. – (18) **8**: tetr. ArEtPrGeX. Schwartz and Lewinsohn 1931 O. $M_4[GeO_4]$: Goldschmidt 1931 X. $M''_2[GeO]_4$. Schutz 1936 X. GeH_4: Steward and Nielsen 1935 S. $Ge(CH_3)_4$: Brockway and Jenkins 1936 E $GeCl_4$: Brockway 1935 E. $GeBr_4$. Rouault 1938*a* E. GeI_4. Hassel and Sandbo 1938 X. GeS_2· Zachariasen 1936 X. Cf. Goldschmidt 1926*c*,*d* X.

Ge IV 6: – (18) **12**: oct. $M_2[GeF_6]$: Wyckoff and Muller 1927 X; Hoard and Vincent 1939 X.

Tin. Sn IV 4: – (18) **8**: tetr Grey tin: Bijl and Kolkmeijer 1919 X. MeEtPrSnX: Pope and Peachey 1900 O $Sn(CH_3)_4$: Brockway and Jenkins 1936 E. $Sn(C_6H_5)_4$: Giacomello 1938 X. $SnCl_4$· Brockway and Wall 1934 E. SnI_4: Hassel and Sandbo 1938 X.

Sn IV 6· – (18) **12**: oct. $M_2[SnCl_6]$: Dickinson 1922*a* X, Wyckoff and Corey 1929 X. $M_2[SnBr_6]$: Ketelaar, Rietdijk and v. Staveren 1937 X. $M_2[SnI_6]$: Werker 1939 X. $K_2[Sn(OH)_6]$. Wyckoff 1928*a* X.

Sn II 2: – (18) (2) **4**: triangular. $SnCl_2$, E in vapour: Sutton and Lister 1940 unpublished.

Lead. Pb IV 4. – (18) **8**: tetr. $Pb(CH_3)_4$. Brockway and Jenkins 1936 E. $Pb(C_6H_5)_4$: George 1927 X; Giacomello 1938 X.

Pb IV 6. – (18) **12**: oct. $M_2[PbCl_6]$: Wyckoff and Dennis 1926 X; Engel 1935 X. $Na_2[Pb(OH)_6]$: Bellucci and Parravano 1906 Cr, Zocher 1920 Cr

Pb II 2: – (18) (2) **4**· triangular $PbCl_2$, $PbBr_2$: Powell and Tasker 1937 X, Sutton and Lister E in vapour 1940 unpublished.

Pb II 4: – (18) (2) **8**: planar. $K_2[Pb(C_2O_4)_2]$, plumbous benzoylacetonate, etc.: Cox, Shorter and Wardlaw 1937 X.

Titanium. Ti IV 4: – (8) **8**: tetr. $TiCl_4$: Wierl 1930 E. $TiBr_4$, TiI_4: Hassel and Kringstad 1932 X.

Ti IV 6: – (8) **12**. oct. $M_2[TiF_6]$: Hassel and Salvesen 1927 X, Zambonini 1930 X. $Cs_2[TiCl_6]$: Engel 1935 X. Catechol complex $M_2[Ti(cat)_3]$: Rosenheim, Reibmann and Schendel 1931 O.

Zirconium. Zr IV 6: (8) **12**: oct. $Zn[ZrF_6]$, $6H_2O$: Hassel and Salvesen 1927 X. $Rb_2[ZrCl_6]$: Engel 1935 X.

Zr IV 7: – (8) **14**: $K_3[ZrF_7]$: Hampson and Pauling 1938 X (correcting Hassel and Mark 1924 X).

Hafnium. Hf IV 6: – (8) **12**: oct. $(NH_4)_2HfF_6$: v. Hevesy 1927, p. 21.

Hf IV 7: – (8) **14**: $(NH_4)_3HfF_7$: Hassel and Mark 1924 X (same lattice as Zr salt).

GROUP V

Nitrogen. N iii 3: (2) **2**, **6**: pyramidal. NH_3: Wright and Randall 1933 S (I-Red), Watson 1927 D. $N(CH_3)_3$: Brockway and Jenkins 1936 E; Steiger 1930, 1931 D. $NH_2(CH_3)$: Thompson 1939 S. $(CH_2)_6N_4$ (Urotropin)· Dickinson and Raymond 1923 X, Gonell and Mark 1923 X; Wyckoff and Corey 1934 X.

N v 4: (2) **8**: tetr. NH_4 ion: first suggestion of tetrahedral form, van't Hoff 1878; first optical resolution, Pope and Peachey 1899; final proof that it is not a square pyramid, Mills and Warren 1925 O. $[N(CH_3)_4]$hal: Wyckoff 1928*b* X $R_1R_2R_3N \rightarrow O$: Meisenheimer 1908 O. $(CH_3)_3N \rightarrow O$· Lister and Sutton 1939 E.

Phosphorus. P iii 3: – (8) **2**, **6**: pyramidal. White P: Maxwell, Hendricks and Mosley 1935 E. Black P: Hultgren, Gingrich and Warren 1935 X; Gingrich and Hultgren 1935 X $P(CH_3)_3$: Springall and Brockway 1938 E. $P(hal)_3$: Brockway and Wall 1934 E; Gregg *et al.* 1937 E. P_4O_6: Hampson and Stosick 1938 E. $[PO_3]^{--}$: Zachariasen 1931 X.

P v 4: – (8) **8**: tetr. AlP, GaP· Goldschmidt 1927 X; Pauling 1939, p. 165 X. MeEtPrP → O: Meisenheimer and Lichtenstadt 1911 O. $[PCl_4]^+$ in solid PCl_5 and $[PBr_4]^+$ in solid PBr_5: Powell, Clark and Wells 1940 X, Powell and Clark 1940 X. POF_3, $POCl_3$, etc.· Brockway and Beach 1938 E. PSF_3: Stevenson and Russell 1939 E. $M[PO_2F_2]$: Lange 1929 Cr. $(NH_4)_2PO_3F$: Goswami 1937 Cr. $[PO_4]^{---}$: Helmholz 1936 X, cf. Keggin 1934 X. P_4O_{10}: Hampson and Stosick 1938 E. $P_4O_6S_4$: Stosick 1939 E.

P v 5: – (8) **10**: trigonal bipyramid. PF_5· Braune and Pinnow 1937 E; Brockway and Beach 1938 E; Linke and Rohrmann 1937 D. PCl_5. Rouault 1938*b* E; Schomaker (quoted by Pauling 1939) unpublished. Moureu, Magat and Wetroff 1937 S. PF_3Cl_2: Brockway and Beach 1938.

P v 6. – (8) **12**: oct. $[PCl_6]^-$ in solid PCl_5: Powell, Clark and Wells 1940 X.

Arsenic. As iii 3: – (18) **2**, **6**. pyram. As_4: Maxwell, Hendricks and Mosley 1935 E. Solid As: Bradley 1924 X. $As(CH_3)_3$: Springall and Brockway 1938 E; Gregg *et al* 1937 E; Yost and Sherborne 1934 S. $AsBr_3$: Braekken 1935 X. $M_3[AsO_3]$: Zachariasen 1931*a* X.

As v 4· – (18) **8**: tetr MeEtAsS-C_6H_4-COOH Mills and Raper 1925 O $Et_3As \rightarrow$ Cu—I. Wells 1936*a* X. $(Me_3As)_2PdBr_2$. Wells 1938*a* X. GaAs: Goldschmidt 1927 X. Ag_3AsO_4. Wyckoff 1925 X.

As v 6: – (18) **12**: oct. Catechol complexes, $M[As(cat)_3]$. Rosenheim and Plato 1925 O.

Antimony. Sb iii 3. – (18) **2**, **6**: pyram. $SbCl_3$, $SbBr_3$, SbI_3: Gregg *et al.* 1937 E. $SbBr_3$: Braekken 1935 X. Sb_4O_6. Bozorth 1923 X, Buerger and Hendricks 1937 E. $M_3[SbO_3]$: Zachariasen 1931*a* X.

Sb v 4: – (18) **8**: tetr. GaSb: Goldschmidt 1927 X.

Sb v 5: – (18) **10**: trg. bipyr. $(CH_3)_3Sb(hal)_2$: Wells 1938*b* X.

Sb v 6: – (18) **12**. oct. $M[Sb(OH)_6]$: Beintema 1936 X. $Na[SbF_6]$: Schrewelius 1938 X.

Sb ? 6: valency obscure. oct. $M_2[SbCl_6]$ and $M_2[SbBr_6]$: Jensen 1937 X.

Bismuth. Bi iii 3. – (18) **2**, **6**: pyram. $BiCl_3$, $BiBr_3$: Skinner and Sutton 1940*b* E.

Bi iii 6: – (18) (2) **12**. oct. Bi_2O_3 γ-form: Sillen 1938 X.

Vanadium. V v 4: – (8) **8**: *tetr.* $VOCl_3$: Palmer 1938 E. $Y[VO_4]$: *Broch 1933 X.*

V iii 6: – (10) **12**: oct. $(NH_4)_3VF_6$: Passerini and Pirani 1932*a* X.

Niobium. Nb v 4: – (8) **8**: tetr. $Y[NbO_4]$: Barth 1926 X.

Nb v 5: – (8) **10**: trig. bipyramid. $NbCl_5$, $NbBr_5$: Skinner and Sutton 1940*a* E.

Nb v 7: – (8) **14**. $K_2[NbF_7]$: Hoard 1939 X.

Tantalum: references are identical with those for Nb.

GROUP VI

Oxygen. O II 2: (2) 4, 4: triang. H_2O: Mecke and Baumann 1932 S; Sänger 1930 D. $(CH_3)_2O$: Sänger and Steiger 1929 D, Steiger 1930 D. $(NH_4)_2S_2O_8$: Zachariasen and Mooney 1934 X.

But O II 2 seems to be sometimes linear: $[P_2O_7]^{iv}$: Levi and Peyronel 1936 X. $[Si_2O_7]$ ion. Gossner & Mussgnug 1929; Zachariasen 1930*b*. Cristobalite: Wyckoff 1925*a*.

O IV 3 (2) 2, 6: prob. pyram. in $Tl_4(O\ Alk)_4$: Sidgwick and Sutton 1930.

O VI 4: (2) 8. tetr. $Be_4O(OCO.CH_3)_6$: Morgan and Bragg 1923 X. $Zn_4O(OCO.CH_3)_6$: Wyart 1926 X. Also in all oxides with zinc blende or wurtzite structures.

Sulphur. S II 2: – (8) 4, 4: triang. Rhombic S: Warren and Burwell 1935 X. H_2S: Pauling 1939, p. 79 S, Zahn and Mills 1928 D. $(CH_3)_2S$ Brockway and Jenkins 1936 E, Hunter and Partington 1932 D. Trithionates: Zachariasen 1934*b* X.

S IV 3 – (8) 2, 6: pyram. $MeEtS\cdot CH_2\cdot CO_2H[X]$: Pope and Peachey 1900*b* O. Sulphinic esters: Phillips 1925 O. Sulphoxides: Harrison, Kenyon and Phillips 1926 O $[SO_3]''$: Zachariasen and Buckley 1931 X.

S VI 4: – (8) 8: tetr. ZnS, CdS, etc.: Pauling 1939 p. 165 X. $M_2[SO_4]$: Taylor and Boyer 1927 X; Zachariasen and Ziegler 1932 X, Beevers and Lipson 1932 X. $SO_2(hal)_2$: Palmer 1938 E, Stevenson and Russell 1939 E. $(CH_3)_2SO_2$: Lister and Sutton 1939 E

S VI 6: – (8) 12 oct. SF_6: Brockway and Pauling 1933 E.

Selenium. Se II 2: – (18) 4, 4: triang. $(CH_3)_2Se$: Donzelot 1936 S. $(C_6H_5)_2Se$: Bergmann, Engel and Sandor 1930 D.

Se IV 3. – (18) 2, 6: pyram MePhSe $CH_2.CO_2H[X]$: Pope and Neville 1902 O $M_2[SeO_3]$: Zachariasen 1931 X.

Se VI 4. – (18) 8: tetr. ZnSe, CdSe, MgSe, BeSe, Pauling 1939, p 165 X. M_2SeO_4 isomorphous with M_2SO_4: Pellini 1909, 1910 Cr; Patry 1936 X

Se VI 6 (18) 12: oct. SeF_6 Brockway and Pauling 1933 E.

Se IV 6: (18) (2) 12: oct. $M_2[SeCl_6]$: Engel 1934 X. $M_2[SeBr_6]$: Carozzi 1924 Cr, Sieg 1932 X; Hoard and Dickinson 1933 X.

Tellurium. Te II 2: – (18) 4, 4: triang. $(C_6H_5)_2Te$. Bergmann, Engel and Sándor 1930 D.

Te IV 3: – (18) 2, 6: pyram. PhTolMeTe[*X*]: Lowry and Gilbert 1929 O.

Te VI 4: – (18) 8: tetr. ZnTe, HgTe, CdTe. Zachariasen 1926*a* X. M_2TeO_4 isomorphous with M_2SO_4: *Pellini 1909, 1910 Cr, Hammel 1936 X; Patry 1936 X.*

Te VI 6. – (18) 12. oct. TeF_6: Brockway and Pauling 1933 E. $Te(OH)_6$: Pauling 1935 X.

Te IV 6: – (18) (2) 12: oct. $M_2[TeCl_6]$: Natta and Pirani 1932 X.

Chromium. Cr VI 4. – (8) 8: tetr. $M_2[CrO_4]$: Zachariasen and Ziegler 1931 X, Clouse 1932 X; Miller 1936, 1938 X. CrO_2Cl_2: Palmer 1938 E.

Cr III 6. – (11) 12: oct. $[Cr(en)_2Cl_2]X$: Werner 1911*d* O. $K_3[Cr(C_2O_4)_3]$: Werner 1912*d* O. $[Cr(OH_2)_6]Cl_3$: Andress and Carpenter 1934 X. $(NH_4)_3CrF_6$: Passerini and Pirani 1932*b* X. $K_3[Cr(CN)_6]$: Gottfried and Nagelschmidt 1930 X.

Cr 0 6: – (14) 12: oct. $Cr(CO)_6$: Rudorff and Hofmann 1935 X; Brockway, Ewens and Lister 1938 E.

Molybdenum. Mo VI 4: – (8) 8. tetr. $Ag_2[MoO_4]$: Wyckoff 1922*b* X; Dickinson 1920 X.

Mo VI 6. – (8) 12: oct. MoF_6: Braune and Pinnow 1937 E. MoO_6 groups in hetero-poly salts: Keggin 1934 X; Santos 1935 X. $(NH_4)_3[MoO_3F_3]$: Pauling 1924 X.

Mo V 5: – (9) 10: trig. bipyramid. $MoCl_5$: Ewens and Lister 1938 E.

Mo IV 8: – (10) 16: $K_4[Mo(CN)_8]$: Hoard and Nordsieck 1939 X.

Mo III 6: – (11) **12**: oct. $R_3[MoCl_5(OH_2)]$: Carobbi 1928 Cr.

Mo 0 6: – (14) **12**: oct. $Mo(CO)_6$: Rüdorff and Hofmann 1935 X; Brockway, Ewens and Lister 1938 E.

Tungsten. W VI 4: – (8) **8**: tetr. $CaWO_4$: Barth 1926 X.

W VI 6: – (8) **12**: oct. WF_6: Braune and Pinnow 1937 E. WCl_6: Ewens and Lister 1938 E. WO_6 groups in hetero-poly acids: Keggin 1934 X; Santos 1935 X.

W III 6: – (11) **12**: oct. $R_3[W_2Cl_9]$· Brosset 1935, 1936 X.

W 0 6: – (14) **12**: oct. $W(CO)_6$: Rudorff and Hofmann 1935 X: Brockway, Ewens and Lister 1938 E.

Uranium. U VI 6: – (8) **12**: UF_6: Braune and Pinnow 1937 E.

GROUP VII

Fluorine. Polymerized HF, gas and solid, has bent chains F··H··F H··F· H with F angle *ca* 140°: Bauer, Beach and Simons 1939 E; Gunther, Holz and Strunz 1939 E.

Chlorine. ClO_2. 23/3 valency electrons: triang. Sutton and Brockway 1935 E.

Cl III 2: – (8) **4**, **4**: triang. $M[ClO_2]$: Levi and Scherillo 1931 X Cl in $(PdCl_2)_n$. Wells 1938 X.

Cl V 3: – (8) **2**, **6**: pyramid $M[ClO_3]$: Zachariasen 1931 *b* X.

Cl VII 4. – (8) **8**: tetr. $M[ClO_4]$· Bussem and Hermann 1928 X, Zachariasen 1930 X; Gottfried and Schusterius 1933 X

Bromine. Br III 2. – (18) **4**, **4**· triang. $(Alk_2AuBr)_2$· Burawoy, Gibson, Hampson and Powell 1937 X

Br V 3: – (18) **2**, **6**· pyramid. $M[BrO_3]$· Zachariasen 1931 *a* X.

Iodine. I V 3. – (18) **2**, **6**. pyramidal. $(Me_3As \rightarrow Cu—I)_4$· Mann, Purdie and Wells 1936 X, Wells 1937 X.

I VII 4: – (18) **8**. tetr. $M[IO_4]$. Kirkpatrick and Dickinson 1926 X; Hylleraas 1926 X.

I VII 6· (18) **12**. oct. $(NH_4)_2H_3(IO_6)$: Helmholz 1937 X.

• I III 2: (18) (2) **4**, **4**: linear $Cs[ICl_2]$: Wyckoff 1920 X; Hassel 1931 X. $NH_4[I_3]$: Mooney 1935 *a* X. M[Cl-I-Br]: Mooney 1935 *b* X, 1938 *a* X, 1939 X.

I V 5. – (18) (2) **10**: IF_5: positions uncertain: Braune and Pinnow 1937 E.

I III 4: – (18) (4) **8**: planar. $M[ICl_4]$: Mooney 1938 *b* X.

Manganese. Mn VII 4· – (8) **8**: tetr. $M[MnO_4]$: Basche and Mark 1926 X. Mooney 1931 X.

Mn VI 4. (9) **8**. $M_2[MnO_4]$. Cox, Shorter, Wardlaw and Way 1937 X.

Mn II 4 – (13) **8**: (A) planar, (B) tetr. A. Planar: py_2MnCl_2: Cox, Shorter, Wardlaw and Way 1937 X. K_2SO_4, $[Mn(OH_2)_4]SO_4$: Anspach 1939 X. B Tetr MnSe: Baroni 1938 X.

Mn II 6: – (13) **12**: oct. $[Mn(NH_3]_6]X_2$: Bodtker-Naess and Hassel 1933 *a* X. $[Mn(OH)_6]SiF_6$: Hassel 1927 X; Hassel and Salvesen 1927 X.

Rhenium. Re VII 4: – (8) **8**· tetr. $M[ReO_4]$. Broch 1929 X; Jaeger and Beintema 1933 X; Pitzer 1935 X; Beintema 1937 X.

Re VI 4: – (9) **8**: tetr. $M_2[ReO_4]$ Noddacks 1933 Cr.

GROUP VIII

Iron. Fe IV 7: – (12) **14**. $Fe_2(CO)_9$: Powell and Ewens 1939 X.

Fe III 6: – (13) **12**: oct. $M_3[Fe(C_2O_4)_3]$· Thomas 1921 X. $K_3[Fe(CN)_6]$· Gottfried and Nagelschmidt 1930 X, Keggin and Miles 1936 X. $(NH_4)_3[FeF_6]$: Pauling 1924 X.

Fe II 6: – (14) 12: oct. [Tri-dipyridyl Fe]X_2: Werner 1912*b* O. $[Fe(NH_3)_6]X_2$. Bödtker-Naess and Hassel 1933*a* X. $M_4[Fe(CN)_6]$: Brill and Mark 1928 X; Keggin and Miles 1936 X. $M_2'M''[Fe(NO_2)_6]$: Cambi and Ferrari 1935 X.

Fe 0 5: – (16) 10: trig. bipyramid. $Fe(CO)_5$: Ewens and Lister 1939 E.

Fe –II 4: – (18) 8: tetr. $Fe(CO)_2(NO)_2$: Brockway and Anderson 1937 E. $Fe(CO)_2(COH)_2$: Ewens and Lister 1939 E.

Cobalt. Co III 6: (14) 12: oct. $[Co(en)_2(NH_3)Cl]X_2$: Werner 1911*a* O. $[Co(en)_2.(NO_2)_2]X$: Werner 1911*b* O. $[Co(en)_2(NO_2)Cl]X$· Werner 1911*c* O. $[Co(en)_3]X_3$: Werner 1911*c* O, 1912*a* O. $M_3[Co(C_2O_4)_3]$: Jaeger 1919 O. $M[Co(NH_3)_2(NO_2)_2(C_2O_4)]$: Thomas 1923 O. $[Co(NH_3)_6]X_3$: Wyckoff and McCutcheon 1927 X; Wyckoff, Hendricks and McCutcheon 1927 X; Hassel 1928 X $Ag[Co(NH_3)_2(NO_2)_4]$: Wells 1936*b* X.

Co II 4. – (15) 8: (A) planar, (B) tetr. A. Planar. $Co(py)_2Cl_2$. Hantzsch 1927 I; Cox, Shorter, Wardlaw and Way 1937 X. B. Tetr. $M_3[CoCl_4](Cl)$. Powell and Wells 1935*a* X; Krishnan and Mookherji 1937 Magnetic.

Co II 6· – (15) 12· oct. $[Co(NH_3)_6]Cl_2$· Stoll 1926 X. $[Co(OH_2)_6]SiF_6$: Hassel 1927 X; Hassel and Salvesen 1927 X.

Co –I 4: – (18) 8: tetr. $Co(CO)_3NO$. Brockway and Anderson 1937 E. $Co(CO)_3COH$ Ewens and Lister 1939 E.

Nickel. Ni II 4: – (16) 8: planar. Glyoxime compounds. Sugden 1932 I. See also Cox, Pinkard, Wardlaw and Webster 1935 X; Cox, Wardlaw and Webster 1935 X.

Ni II 6: – (16) 12· oct. $[Ni(dipy)_3]X_2$: Morgan and Burstall 1931 O. $[Ni(NH_3)_6]X_2$: Wyckoff 1922*a* X, Bodtker-Naess and Hassel 1933*b* X. $[Ni(OH_2)_6]SiF_6$: Hassel 1927 X; Hassel and Salvesen 1927 X. $Ba_2[Ni(NO_2)_6]$: Ferrari and Curti 1933 X.

Ni O 4: – (18) 8: tetr. $Ni(CO)_4$· Brockway and Cross 1935 E; Bailey and Gordon 1938 S.

Ruthenium. Ru IV 6· – (12) 12: oct. $K_2[RuCl_6]$· Howe 1904 Cr; Howe and Haynes 1925 Cr.

Ru III 6· – (13) 12· oct. Acetonylacetonate RuA_3. Barbieri 1914 Cr.

Ru II 6: – (14) 12. oct. $K[Ru(C_2O_4)_2(NO)py]$: Charronat 1924 O. $[Ru(dipy)_3]X_2$: Burstall 1936 O. See also Keggin and Miles 1936 X.

Rhodium. Rh III 6· – (14) 12. oct. $[Rh(en)_3]X_3$: Werner 1912*c* O. Cf. Jaeger and Blumendal 1928 O, Mann 1933 O. $M_3[Rh(C_2O_4)_3]$: Werner 1914 O. $M_3[Rh(NO_2)_6]$: Ferrari and Colla 1933*b* X. $[Rh(NH_3)_3Cl]^0$: West 1935*b* X.

Palladium. Pd IV 6: – (14) 12: oct. $Rb_2[PdBr_6]$: Scherrer and Stoll 1922 X.

Pd II 4: – (16) 8: planar. PdO: Lunde 1927 X. PdS: Gaskell 1937 X. $PdCl_2$: Wells 1938*c* X. $(Me_3As)_2Pd_2Br_4$: Mann and Wells 1938 X, Wells 1938*a* X. $(Me_2S)_2PdCl_2$: Cox, Saenger and Wardlaw 1934 X. Acetylacetonate PdA_2: Barbieri 1914 Cr. $[Pd(NH_3)_4]Cl_2$: Cox and Preston 1933 X; Dickinson 1934 X. $K_2[PdCl_4]$: Dickinson 1922*b* X; Cox and Preston 1933 X, Theilacker 1937 X. Complex dithio-oxalate: Cox, Wardlaw and Webster 1935 X. Proof of planar structure by optical activity: Lidstone and Mills 1939 O. Cf also in general Jensen 1939.

Osmium. Os VIII 4: – (8) 8: tetr. OsO_4: Brockway 1936 E; Langseth and Qviller 1934 S. $K[OsO_3N]$: Jaeger and Zanstra 1932 X

Os VI 6: – (10) 12: oct. $M_2[OsO_2Cl_4]$: Verhulst 1933 X. $M_2[OsNCl_5]$: Hoard and Grenko 1934 X.

Os IV 6: – (12) 12: oct. $M_2[OsCl_6]$, $M_2[OsBr_6]$: McCullough 1936 X.

Iridium. Ir IV 6· – (13) 12· oct. $[Ir(py)_2Cl_4]^0$: Delepine 1922 Cr, I. $K_2[Ir(C_2O_4)_3]$: Delepine 1917 O.

Ir III 6: – (14) 12: oct. $M_3[Ir(C_2O_4)_3]$: Delepine 1914 O. $[Ir(en)_3]X_3$: Werner and Smirnoff 1920 O. $M_3[Ir(CN)_6]$: Gottfried and Nagelschmidt 1930 X. $M_3[Ir(NO_2)_6]$: Ferrari and Colla 1933*a* X.

Platinum. Pt IV 4: – (14) 8: tetr. $Pt(CH_3)_3Cl$: Cox and Webster 1935*b* X.

Pt IV 6: – (14) 12: oct. $[Pt(en)_3]X_4$: Werner 1917 O. $M_2[PtCl_6]$: Wyckoff and Posnjak 1921 X; Scherrer and Stoll 1922 X.

Pt II 4: – (16) 8: planar: first suggested by Werner 1893 I· supported by dipole moments, Jensen 1936 D: for answer to objections see Jensen 1939. PtS: Bannister 1932 X. $K_2[PtCl_4]$: Dickinson 1922*b* X. $K_2[Pt(CN)_4]$: Brasseur and de Rassenfosse 1938 X. $[Pt(NH_3)_4]Cl_2$: Cox 1932 X; Dickinson 1934 X. Stereochemical proof of plane structure. Mills and Quibell 1935 O

REFERENCES

Andress, K. R. and Carpenter, C. 1934 *Z. Kristallogr.* **87**, 446.
Anspach, H. 1939 *Z. Kristallogr* **101**, 39.
Bailey, C. R. and Gordon, R. R. 1938 *J. Chem. Phys.* **6**, 225.
Bailey, C. R., Hall, J. B. and Thompson, J W 1937 *Proc. Roy. Soc.* A, **161**, 107.
Bannister, F. A. 1932 *Min. Mag.* **23**, 188.
Barbieri, G. A. 1914 *R C. Accad. Lincei*, [5], **23**, 334.
Baroni, A. 1938 *Z. Kristallogr.* **99**, 336.
Barth, T. 1926 *Norsk. geol. Tidskr.* **9**, 24.
Basche, W. and Mark, H. 1926 *Z. Kristallogr.* **64**, 1.
Bauer, S. H. 1937 *J. Amer. Chem. Soc.* **59**, 1096.
Bauer, S H., Beach, J. Y. and Simons, J. H. 1939 *J. Amer. Chem. Soc* **61**, 19.
Beevers, C. A. and Lipson, H 1932 *Z. Kristallogr.* **82**, 297.
Beintema, J. 1936 *Proc. Acad. Sci. Amst.* **39**, 241, 652.
Bellucci, I. and Parravano, N. 1906 *Z. anorg. Chem.* **50**, 101.
Bergmann, E., Engel, L. and Sándor, S. 1930 *Z. Phys. Chem.* B, **10**, 397.
Bijl, A. J. and Kolkmeijer, N. H. 1919 *Proc. Acad Sci Amst.* **21**, 494
Biltz, W. and Voigt, A. 1924 *Z. anorg. Chem.* **133**, 297.
Bloch, R. and Möller, H. 1931 *Z. Phys. Chem.* A, **152**, 245.
Bodtker-Naess, G. and Hassel, O. 1933*a* *Z. Phys. Chem.* B, **22**, 471
— — 1933*b* *Z. anorg. Chem.* **211**, 21.
Boeseken, J. and Meulenhoff, J. 1924 *Proc. Acad. Sci. Amst.* **27**, 171.
Bozorth, R M. 1923 *J. Amer. Chem. Soc* **45**, 1621.
Bradley, A. J. 1924 *Phil. Mag.* [vi] **47**, 657.
Braekken, H. 1935 *K. Norske Vidensk. Selsk. Forh.* **8**, no. 10. (*Chem Zbl.* **36**, 1, 961.)
Braekken, H. and Scholten, W. 1934 *Z. Kristallogr.* **89**, 448.
Bragg, W. H. 1921 *Proc. Phys. Soc.* **33**, 304.
Bragg, W. H. and W. L. 1913 *Nature, Lond.*, **91**, 557; *Proc. Roy. Soc.* A, **89**, 277.
Bragg, W. H. and Gibbs, R. E. 1925 *Proc. Roy. Soc.* A, **109**, 405.
Brandenberger, E. 1932 *Schweiz. Min. petr. Mitt.* **12**, 243.
Brasseur, H. and de Rassenfosse, A. 1938 *Bull. Soc. Fr. Min.* **61**, 129.
Braune, H. and Knoke, S. 1933 *Z. Phys. Chem.* B, **23**, 163.
Braune, H. and Linke, R. 1935 *Z. Phys. Chem.* B, **31**, 12.
Braune, H. and Pinnow, P. 1937 *Z. Phys. Chem.* B, **35**, 239.
Brill, R. and Mark, H. 1928 *Z. Phys. Chem* **133**, 443.
Broch, E. 1929 *Z. Phys. Chem.* B, **6**, 22.
— 1933 *Z. Phys. Chem.* B, **20**, 345.

Brockway, L. O. 1935 *J. Amer. Chem. Soc.* **57**, 958.
— 1936 *Rev. Mod. Phys.* **8**, 260.
Brockway, L. O. and Anderson, J. S. 1937 *Trans. Faraday Soc.* **33**, 1233.
Brockway, L. O. and Beach, J. Y. 1938 *J. Amer. Chem. Soc.* **60**, 1836.
Brockway, L. O. and Cross, P. C. 1935 *J. Chem. Phys.* **3**, 828.
Brockway, L. O., Ewens, R. V. G. and Lister, M. W. 1938 *Trans. Faraday Soc.* **34**, 1350.
Brockway, L. O. and Jenkins, H. O. 1936 *J. Amer. Chem. Soc.* **58**, 2036.
Brockway, L. O. and Pauling, L. 1933 *Proc. Nat. Acad. Sci., Wash.*, **19**, 68.
Brockway, L. O. and Wall, F. T. 1934 *J. Amer. Chem. Soc.* **56**, 2373.
Brosset, C. 1935 *Nature, Lond.*, **135**, 874.
— 1936 *Ark. Kemi Min. Geol.* **12 A**, no. 4.
Buerger, M. J. and Hendricks, S. B. 1937 *J. Chem. Phys.* **5**, 600.
Büssem, W., Fischer, H. and Gruner, E. 1935 *Naturwissenschaften*, **23**, 740.
Bussem, W. and Herrmann, K. 1928 *Z. Kristallogr.* **67**, 405.
Burawoy, A., Gibson, C. S., Hampson, G. C. and Powell, H. M. 1937 *J. Chem. Soc.* p. 1690.
Burgess, H. and Lowry, T. M. 1924 *J. Chem. Soc.* **125**, 2081.
Burstall, F. H. 1936 *J. Chem. Soc.* p. 173.
Cambi, L. and Ferrari, A. 1935 *Gaz. chim. ital.* **65**, 1162.
Carobbi, G. 1928 *Gaz. chim. ital.* **58**, 35.
Carozzi, E. 1924 *Gaz. chim. ital.* **54**, 556.
Charronat, R. 1924 *C.R. Acad. Sci., Paris*, **178**, 1423
Chrobak, L. 1934 *Z. Kristallogr.* **88**, 35.
Clouse, J. H. 1932 *Z. Kristallogr.* **83**, 161.
Corey, R. B. and Wyckoff, R. W. G. 1934 *Z. Kristallogr.* **87**, 264.
Cox, E. G. 1932 *J. Chem. Soc.* p. 1912.
Cox, E. G., Pinkard, F. W., Wardlaw, W. and Webster, K. C. 1935 *J. Chem. Soc.* p. 459.
Cox, E. G. and Preston, G. H. 1933 *J. Chem. Soc.* p. 1089.
Cox, E. G., Saenger, H. and Wardlaw, W. 1934 *J. Chem. Soc.* p. 182.
Cox, E. G., Sharratt, E., Wardlaw, W. and Webster, K. C. 1936 *J. Chem. Soc.* p. 129.
Cox, E. G., Shorter, A. J. and Wardlaw, W. 1937 *Nature, Lond.*, **139**, 71
— — — 1938 *J. Chem. Soc.* p. 1886.
Cox, E. G., Shorter, A. J., Wardlaw, W. and Way, W. J. R. 1937 *J. Chem. Soc.* p. 1556.
Cox, E. G., Wardlaw, W. and Webster, K. C. 1935 *J. Chem. Soc.* p. 1475
— — — 1936 *J. Chem. Soc.* p. 775.
Cox, E. G. and Webster, K. C. 1935*a* *J. Chem. Soc.* p. 731.
— 1935*b* *Z. Kristallogr.* **90**, 561.
— — 1936 *J. Chem. Soc.* p. 1635.
Crenshaw, J. L., Cope, A. C., Finkelstein, N. and Rogan, R. 1938 *J. Amer. Chem. Soc.* **60**, 2308.
Curran, W. J. and Wenzke, H. H. 1935 *J. Amer. Chem. Soc.* **57**, 2162.
Degard, C., Piérard, J. and van der Grinten, W. 1935 *Nature, Lond.*, **136**, 142.
Delepine, M. 1914 *C.R. Acad. Sci., Paris*, **159**, 239.
— 1917 *Bull. Soc. chim. Paris*, [iv] **21**, 157.
— 1922 *C.R. Acad. Sci., Paris*, **175**, 1211.
Dickinson, B. N. 1934 *Z. Kristallogr.* **88**, 281.
Dickinson, R. G. 1920 *J. Amer. Chem. Soc.* **42**, 85.
— 1922*a* *J. Amer. Chem. Soc.* **44**, 276.
— 1922*b* *J. Amer. Chem. Soc.* **44**, 774, 2404.

Dickinson, R. G. and Raymond, A. L. 1923 *J. Amer. Chem. Soc.* **45**, 22.

Donzelot, P. 1936 *C.R. Acad. Sci., Paris*, **203**, 1069.

Dothie, H. J., Llewellyn, F. J., Wardlaw, W. and Welch, A. J. E. 1939 *J. Chem. Soc.* p. 426.

Ehrenberg, W. 1926 *Z. Kristallogr.* **63**, 320.

Elliott, N. and Pauling, L. 1938 *J. Amer. Chem. Soc.* **60**, 1846.

Engel, G. 1934 *Zbl. Min. Geol. Paläont.* A, 285.

— 1935 *Z. Kristallogr.* **90**, 341.

Ewens, R. V. G. and Lister, M. W. 1938 *Trans. Faraday Soc.* **34**, 1358.

— — 1939 *Trans. Faraday Soc.* **35**, 681

Ferrari, A. and Colla, C. 1933*a* *Gaz. chim. ital.* **63**, 507.

— — 1933*b* *R.C. Accad. Lincei*, [VI], **18**, 45.

Ferrari, A. and Curti, R 1933 *Gaz. chim. ital.* **63**, 499.

Finbak, C. and Hassel, O. 1937 *Z. Phys. Chem.* B, **36**, 301.

Gaskell, T F. 1937 *Z. Kristallogr.* **96**, 203.

George, W. H. 1927 *Proc. Roy. Soc.* A, **113**, 585.

Giacomello, G. 1938 *Gaz. chim. ital.* **68**, 422.

Gibbs, R. E. 1926 *Proc. Roy. Soc.* A, **110**, 443.

Gibson, C. S. 1939 *Proc. Roy. Soc.* A, **173**, 160.

Gingrich, N. S and Hultgren, R. 1935 *Phys. Rev.* ii, **47**, 808.

Goldschmidt, V. M. 1926*a* *Geochem. Vert. Oslo*, II.

— 1926*b* *Geochem Vert Oslo*, VII.

— 1927 *Geochem. Vert. Oslo*, VIII.

— 1931 *Nachr. Ges. Wiss. Göttingen*, 184.

Gonell, H. W. and Mark, H. 1923 *Z. Phys Chem.* A, **107**, 181.

Gossner, B. and Mussgnug, F. 1929 *Zbl. Min. Geol. Paläont.* A, 1.

Goswami, H. C. 1937 *J. Ind Chem. Soc* **14**, 660.

Gottfried, C. and Nagelschmidt, J. G. 1930 *Z. Kristallogr.* **73**, 357.

Gottfried, C. and Schusterius, C 1933 *Z. Kristallogr* **84**, 65.

Gregg, A. H., Hampson, G. C., Jenkins, G. I., Jones, P. L. F. and Sutton, L. E. 1937 *Trans. Faraday Soc* **33**, 852.

Grimm, H. G. and Sommerfeld, A. 1926 *Z. Phys.* **36**, 44.

Gunther, P., Holm, K. and Strunz, H. 1939 *Z. Phys. Chem* B, **43**, 229

Hammel, F. 1936 *C.R. Acad. Sci., Paris*, **202**, 57, 2147.

Hampson, G. C. 1934 *Trans Faraday Soc.* **30**, 877.

Hampson, G C. and Pauling, L. 1938 *J. Amer. Chem. Soc.* **60**, 2702.

Hampson, G. C. and Stosick, A. J. 1938 *J. Amer. Chem Soc.* **60**, 1814.

Hantzsch, A. 1927 *Z. anorg. Chem.* **159**, 278.

Harker, D. 1936 *Z. Kristallogr.* **93**, 136.

Harrison, P. W B., Kenyon, J. and Phillips, H. 1926 *J. Chem Soc.* p. 2079.

Hassel, O. 1927 *Z. Phys. Chem.* A, **126**, 118.

— 1928 *Norsk. geol. Tidskr.* **10**, 92

— 1931 *Tidskr. Kemi Bergvesen*, **11**, 92.

Hassel, O. and Kringstad, H. 1931 *Z. Phys. Chem.* B, **13**, 1.

— — 1932 *Z. Phys. Chem.* B, **15**, 274

Hassel, O. and Mark, H. 1924 *Z. Phys.* **27**, 89.

Hassel, O. and Salvesen, J. R. 1927 *Z. Phys Chem* A, **128**, 345.

Hassel, O. and Sandbo, A. 1938 *Z. Phys. Chem.* B, **41**, 75.

Havighurst, R. J. 1925 *Amer. J. Sci.* **10**, 15.

— 1926 *J. Amer. Chem. Soc.* **48**, 2113.

Hein, F. and Regler, H. 1936 *Ber. dtsch. chem. Ges.* **69**, 1692.

Helmholz, L. 1935 *J. Chem. Phys.* **3**, 740.

Helmholz, L. 1936 *J. Chem. Phys.* **4**, 316.
— 1937 *J. Amer. Chem. Soc.* **59**, 2036.
de Hemptinne, M. and Wouters, J. 1936 *Nature, Lond.*, **138**, 884.
Hendricks, S. B. and Dickinson, R. G. 1927 *J. Amer. Chem. Soc.* **49**, 2149.
v. Hevesy, G. 1927 *Das Hafnium* (Berlin · Springer).
Hoard, J. L. 1933 *Z. Kristallogr.* **84**, 231.
— 1939 *J. Amer. Chem. Soc.* **61**, 1252.
Hoard, J. L. and Blair, V. 1935 *J. Amer. Chem. Soc.* **57**, 1985.
Hoard, J. L. and Dickinson, B. N. 1933 *Z. Kristallogr.* **84**, 436.
Hoard, J. L. and Goldstein, L. 1935*a* *J. Chem. Phys.* **3**, 117.
— — 1935*b* *J. Chem. Phys.* **3**, 199.
— — 1935*c* *J. Chem. Phys.* **3**, 645.
Hoard, J. L. and Grenko, J. D. 1934 *Z. Kristallogr.* **87**, 100.
Hoard, J. L. and Nordsieck, H. H. 1939 *J. Amer. Chem. Soc.* **61**, 2853.
Hoard, J. L. and Vincent, W. B. 1939 *J. Amer. Chem. Soc.* **61**, 2849.
van't Hoff, J. H. 1874 *Voorstel tot Uitbreiding der Structoor Formules in de Ruimte.* Utrecht, Sept. 1874
— 1878 *Ansichten u. d. organ. Chemie*, p. 80.
Huggins, M. L. and Magill, P. L. 1927 *J. Amer. Chem. Soc.* **49**, 2357.
Hultgren, R. 1934 *Z. Kristallogr.* **88**, 233
Hultgren, R., Gingrich, N. S. and Warren, B. E. 1935 *J. Chem. Phys.* **3**, 351.
Hunter, E. C. E. and Partington, J. R. 1932 *J. Chem. Soc.* p. 2812.
Hylleraas, E. 1925 *Z. Phys.* **36**, 859
— 1926 *Z. Phys.* **39**, 308.
Jaeger, F. M. 1919 *Rec. Trav. chim. Pays-Bas*, **38**, 247, 253.
Jaeger, F. M. and Beintema, J. 1933 *Proc. Acad. Sci. Amst.* **36**, 523.
Jaeger, F. M. and Blumendal, H. B. 1928 *Z. anorg. Chem.* **175**, 161.
Jaeger, F. M. and Zanstra, J. E. 1932 *Proc. Acad. Sci. Amst.* **35**, 610, 787.
Jensen, K. A. 1936 *Z. anorg. Chem.* **229**, 225
— 1937 *Z. anorg. Chem.* **232**, 193
— 1939 *Z. anorg. Chem.* **241**, 115.
Johnson, C. H. and Bryant, S. A. 1934 *J. Chem. Soc.* p. 1783.
Juza, R. and Hahn, H. 1938 *Z. anorg. Chem.* **239**, 285.
Keggin, J. F. 1933 *Nature, Lond.*, **131**, 908
— 1934 *Proc. Roy. Soc.* A, **144**, 75.
Keggin, J. F. and Miles, F. G. 1936 *Nature, Lond.*, **137**, 577.
Ketelaar, J. A. A. 1931 *Z. Kristallogr.* **80**, 190.
— 1935 *Z. Kristallogr.* **92**, 155.
Ketelaar, J. A. A., Rietdijk, A. A. and van Staveren, C. H. 1937 *Rec. Trav. chim. Pays-Bas*, **56**, 907.
Kimball, G. E. 1940 *J. Chem. Phys.* **8**, 188.
Kipping, F. S. 1907 *J. Chem. Soc.* **91**, 209.
— 1908 *J. Chem. Soc.* **93**, 457.
Kirkpatrick, L. M. and Dickinson, R. G. 1926 *J. Amer. Chem. Soc.* **48**, 2327.
Klinkenberg, J. L. and Ketelaar, J. A. A. 1935 *Rec. Trav. chim. Pays-Bas*, **54**, 959.
Kolkmeijer, N. H., Bijvoet, J. M. and Karssen, A. 1924 *Proc. Acad. Sci. Amst.* **27**, 390, 847.
Krishnan, K. S. and Mookherji, A. 1937 *Phys. Rev.* ii, **51**, 774.
Lange, W. 1929 *Ber. dtsch. chem. Ges.* **62**, 786.
Langseth, A. and Qviller, B. 1934 *Z. Phys. Chem.* B, **27**, 79.
Levi, G. R. and Peyronel, G. 1936 *Z. Kristallogr.* **92**, 190.
Levi, G. R. and Scherillo, R. 1931 *Z. Kristallogr.* **76**, 431.

Levy, H. A. and Brockway, L. O. 1937 *J. Amer. Chem. Soc.* **59**, 2085.
Lidstone, A. G. and Mills, W. H. 1939 *J. Chem. Soc.* p. 1754.
Linke, R. and Rohrmann, W. 1937 *Z. Phys. Chem.* B, **35**, 256.
Lister, M. W. and Sutton, L. E. 1939 *Trans Faraday Soc.* **35**, 495.
Lowry, T. M. and Gilbert, F. L. 1929 *J. Chem. Soc.* p. 2867.
Lunde, G. 1927 *Z. anorg. Chem.* **163**, 345.
McCullough, J. D. 1936 *Z. Kristallogr.* **94**, 143.
MacGillavry, C. H. and Bijvoet, J. H. 1936 *Z. Kristallogr* **94**, 249.
Manchot, W. 1922 *Z. anorg. Chem.* **124**, 333.
Mandal K. L. 1939 *Curr. Sci.* **8**, 469.
Mann, F. G. 1933 *J. Chem. Soc.* p 412.
Mann, F. G., Purdie, D. and Wells, A. F. 1936 *J. Chem. Soc.* p. 1503.
Mann, F. G. and Wells, A. F. 1938 *J. Chem. Soc.* p. 702.
Mann, F. G., Wells, A. F. and Purdie, D. 1937 *J. Chem. Soc.* p. 1828.
Mark, H. and Pohland, E. 1925 *Z. Kristallogr.* **62**, 103.
Mark, H. and Steinbach, J. 1926 *Z. Kristallogr.* **64**, 79.
Maxwell, L. R., Hendricks, S. B. and Mosley, V. H. 1935 *J. Chem. Phys.* **3**, 699.
Mecke, R. and Baumann, W. 1932 *Phys. Z.* **33**, 833.
Meisenheimer, J. 1908 *Ber dtsch chem Ges.* **41**, 3966.
Meisenheimer, J. and Lichtenstadt, B. 1911 *Ber. dtsch. chem Ges.* **44**, 356.
Mellor, D. P., Burrows, G. J. and Morris, B. S. 1938 *Nature, Lond.*, **141**, 414.
Menzer, G 1930 *Z. Kristallogr.* **73**, 113
Miller, J. J. 1936 *Z. Kristallogr.* **94**, 131.
— 1938 *Z. Kristallogr.* **99**, 32.
Mills, W. H. and Gotts, R. A 1926 *J. Chem. Soc.* p. 3121.
Mills, W. H. and Quibell, T. H. H. 1935 *J. Chem. Soc.* p. 839.
Mills, W. H. and Raper, R. 1925 *J. Chem. Soc.* **127**, 2479.
Mills, W. H. and Warren, E. H 1925 *J. Chem Soc* **127**, 2507.
Mooney, R. C. L. 1931 *Phys. Rev.* ii, **37**, 1306.
— 1935*a* *Z. Kristallogr.* **90**, 143
— 1935*b* *Phys. Rev.* ii, **47**, 807
— 1938*a* *Z. Kristallogr.* **98**, 324
— 1938*b* *Z. Kristallogr.* **98**, 377.
— 1939 *Z. Kristallogr.* **100**, 519.
Morgan, G. T. and Bragg, W. L. 1923 *Proc. Roy. Soc.* A, **104**, 437.
Morgan, G. T. and Burstall, E H 1931 *J. Chem. Soc.* p. 2213.
Moureu, H, Magat, M. and Wetroff, G 1937 *C.R. Acad. Sci, Paris*, **205**, 276.
Müller, H. and Sack, H. 1930 *Phys. Z.* **31**, 815.
v. Náray-Szabó, S. and Sasvári, K. 1938 *Z Kristallogr* **99**, 27
Natta, G. and Pirani, R. 1932 *R C. Accad. Lincei*, [vi], **16**, 265.
Neogi, P. and Dutt, N. K. 1938 *J. Ind. Chem. Soc.* **15**, 83
Neogi, P. and Mukherjee, G. K 1934 *J Ind. Chem. Soc.* **11**, 225.
Noddack, I. and W. 1933 *Das Rhenium* Leipzig: Barth
Palmer, K. J. 1938 *J. Amer. Chem. Soc.* **60**, 2360.
Palmer, K. J. and Elliott, N. 1938 *J. Amer. Chem. Soc.* **60**, 1852
Passerini, L. and Pirani, R. 1932*a* *Gaz. chim. ital.* **62**, 279.
— — 1932*b* *Gaz. chim. ital* **62**, 289.
Patry, M. 1936 *C.R. Acad. Sci., Paris*, **202**, 1516.
Pauling, L. 1924 *J. Amer. Chem. Soc.* **46**, 2738.
— 1931 *J. Amer. Chem. Soc.* **53**, 1388.
— 1935 *Z. Kristallogr.* **91**, 367.
— 1939 *The Chemical Bond* (quoted by pages). Cornell Univ. Press.

Pellini, G. 1909 *R.C. Accad. Lincei*, [v], **18**, ii, 279.
— 1910 *Gaz. chim. ital* **40**, i, 380: ii, 37.
Pfeiffer, P. and Quehl, K. 1931 *Ber. dtsch. chem. Ges.* **64**, 2667.
— — 1932 *Ber. dtsch. chem. Ges.* **65**, 560.
Phillips, H. 1925 *J. Chem. Soc.* **127**, 2552.
Phillips, R. F. and Powell, H. M. 1939 *Proc. Roy. Soc.* **A**, **173**, 147.
Pirenne, M. H. 1938 *C.R. Acad. Sci*, *Paris*, **206**, 516.
— 1939 *J. Chem. Phys* **7**, 144.
Pitzer, K. S. 1935 *Z Kristallogr.* **92**, 131.
Pope, W J. and Neville, H. A D. 1902 *J. Chem. Soc.* **81**, 1552.
Pope, W. J. and Peachey, S J. 1899 *J. Chem. Soc.* **75**, 1127.
— — 1900 *Proc. Chem. Soc.* **16**, 42, 116; *J. Chem Soc.* **77**, 1072.
Powell, H. M. and Clark, D. 1940 *Nature, Lond.*, **145**, 971.
Powell, H. M., Clark, D. and Wells, A. F. 1940 *Nature, Lond.*, **145**, 149.
Powell, H. M. and Crowfoot, D. M. 1932 *Nature, Lond.*, **130**, 131.
— — 1934 *Z. Kristallogr.* **87**, 370.
Powell, H. M. and Ewens, R V G 1939 *J. Chem. Soc.* p. 286.
Powell, H. M. and Tasker, H. S. 1937 *J. Chem. Soc.* p. 119.
Powell, H. M. and Wells, A. F. 1935 *J. Chem. Soc.* p. 359.
Rosenheim, A. and Plato, W. 1925 *Ber. dtsch. chem. Ges* **58**, 2000.
Rosenheim, A, Reibmann, B. and Schendel, G. 1931 *Z. anorg. Chem.* **196**, 168.
Rouault, M. 1938*a* *C.R. Acad. Sci.*, *Paris*, **206**, 51
— 1938*b* *C R. Acad Sci.*, *Paris*, **207**, 620.
Rüdorff, W and Hofmann, U. 1935 *Z. Phys. Chem.* B, **28**, 351.
Sänger, R. 1930 *Phys. Z.* **31**, 306
Sänger, R. and Steiger, L 1929 *Helv. Phys. Acta*, **2**, 136.
Santos, J. A. 1935 *Proc Roy. Soc* **A**, **150**, 309.
Scherrer, P. and Stoll, P. 1922 *Z. anorg. Chem.* **121**, 319.
Schomaker, V. Unpublished (quoted by Pauling 1939, p. 103 note).
Schrewelius, N. 1938 *Z anorg. Chem.* **238**, 241
Schütz, W. 1936 *Z. Phys. Chem.* B, **31**, 292.
Schwartz, R. and Lewinsohn, M. 1931 *Ber dtsch. chem. Ges.* **64**, 2352.
Sieg, L. 1932 *Z. anorg. Chem.* **207**, 93.
Sillen, L. G. 1938 *Ark. Kemi Min. Geol.* **12**, A, no. 18.
Skinner, H. A. and Sutton, L. E. 1940*a* *Trans. Faraday Soc.* **36**, 668.
— — 1940*b* *Trans. Faraday Soc.* **36**, 681.
Smyth, C. P., Grossman, A. J. and Ginsburg, S. R. 1940 *J. Amer. Chem. Soc.* **62**, 192.
Springall, H D. and Brockway, L. O. 1938 *J. Amer. Chem. Soc.* **60**, 996.
Steiger, O. 1930 *Helv. Phys. Acta*, **3**, 161.
— 1931 *Phys Z.* **32**, 425.
Stevenson, D. P. and Russell, H. 1939 *J. Amer. Chem. Soc.* **61**, 3264.
Stevenson, D. P. and Schomaker, V. 1940 *J. Amer. Chem. Soc.* **62**, 1267.
Steward, W B. and Nielsen, H. H. 1935 *Phys. Rev.* ii, **48**, 861.
Stoll, P. 1926 Diss., Zurich.
Stosick, A. J. 1939 *J. Amer. Chem. Soc.* **61**, 1130.
Sugden, S. 1932 *J. Chem Soc.* p. 246.
Sutton, L. E. and Brockway, L. O. 1935 *J. Amer. Chem. Soc.* **57**, 473.
Sutton, L. E. and Lister, M. W. 1940 Unpublished.
Sutton, L. E. and Sidgwick, N. V. 1930 *J. Chem. Soc.* p 1461.
Taylor, W. and Boyer, T. 1927 *Proc. Manchr Lit. Phil. Soc.* **72**, 125.
Theilacker, W. 1937 *Z. anorg. Chem.* **234**, 161.
Thomas, W. 1921 *J. Chem Soc.* **119**, 1140.

Thomas, W. 1923 *J. Chem. Soc.* **123**, 617.
Thompson, H. W. 1939 *J. Chem. Phys.* **7**, 448.
Treadwell, W. D. 1932 *Helv. chim. Acta*, **15**, 1049.
Ulich, H. and Nespital, W. 1931*a* *Z. angew. Chem* **44**, 750.
— — 1931*b* *Z. Elektrochem.* **37**, 559.
Verhulst, J. 1933 *Bull. Soc. chim. Belge*, **42**, 359.
van Vleck, J. H. and Sherman, A. 1935 *Rev Mod. Phys.* **7**, 167.
Vorländer, D., Hollatz, J. and Fischer, J. 1932 *Ber. dtsch. chem. Ges.* **65**, 535.
Wahl, W. 1927*a* *Ber. dtsch. chem. Ges.* **60**, 399.
— 1927*b* *Acta Soc Sci. Fenn* **4**, (14) 1.
Wardlaw, W. 1940 Unpublished.
Warren, B. E. and Burwell, J. T. 1935 *J. Chem. Phys* **3**, 6.
Warren, B E. and Hill, C. F. 1934 *Z. Kristallogr.* **89**, 481.
Watson, H. E. 1927 *Proc. Roy. Soc.* A, **117**, 43.
Wehrli, M. 1937 *Helv. Phys. Acta*, **10**, 163
Wells, A. F. 1936*a* *Z. Kristallogr.* **94**, 447.
— 1936*b* *Z. Kristallogr.* **95**, 74.
— 1938*a* *Proc. Roy. Soc.* A, **167**, 169
— 1938*b* *Z. Kristallogr.* **99**, 367.
— 1938*c* *Z. Kristallogr.* **100**, 189.
Werker, W. 1939 *Rec Trav chim. Pays-Bas*, **58**, 257.
Werner, A. 1893 *Z. anorg. Chem* **3**, 310.
— 1911*a* *Ber. dtsch. chem. Ges.* **44**, 1887.
— 1911*b* *Ber. dtsch. chem. Ges.* **44**, 2450.
— 1911*c* *Ber. dtsch. chem. Ges.* **44**, 2445, 3272, 3279
— 1911*d* *Ber. dtsch. chem. Ges.* **44**, 3132.
— 1912*a* *Ber. dtsch. chem Ges.* **45**, 121
— 1912*b* *Ber. dtsch. chem. Ges.* **45**, 433.
— 1912*c* *Ber. dtsch. chem. Ges.* **45**, 1228
1912*d* *Ber. dtsch. chem. Ges.* **45**, 3061.
— 1914 *Ber. dtsch. chem. Ges.* **47**, 1954.
— 1917 *Naturf. Ges. Zurich*, **62**, 553
Werner, A. and Smirnoff, A. P. 1920 *Helv chim. Acta* **3**, 472.
West, C. D. 1935*a* *Z. Kristallogr.* **90**, 555.
— 1935*b* *Z. Kristallogr.* **91**, 181.
Wierl, R. 1930 *Phys. Z.* **31**, 366, 1028.
Wright, N. and Randall, H. M. 1933 *Phys. Rev* ii, **44**, 391.
Wyart, J. 1926 *Bull. Soc. Fr. Min* **49**, 148.
Wyckoff, R. W. G. 1920 *J. Amer. Chem. Soc.* **42**, 1100.
— 1922*a* *J. Amer. Chem. Soc.* **44**, 1239, 1260.
— 1922*b* *J. Amer. Chem Soc* **44**, 1994.
— 1925*a* *Z. Kristallogr.* **62**, 189.
— 1925*b* *Amer. J. Sci.* **10**, 107; *Z. Kristallogr.* **62**, 529.
— 1928*a* *Amer. J. Sci.* **15**, 297.
— 1928*b* *Z. Kristallogr.* **67**, 91.
Wyckoff, R. W. G. and Corey, R. B. 1929 *Amer. J. Sci.* **17**, 239; **18**, 437.
— — 1934 *Z. Kristallogr.* **89**, 462.
Wyckoff, R. W G. and Dennis, L. M. 1926 *Amer. J. Sci.* [iv], **12**, 503.
Wyckoff, R. W. G., Hendricks, S. B. and McCutcheon, T. P. 1927 *Amer. J Sci.* **13**, 388.
Wyckoff, R. W. G. and McCutcheon, T. P. 1927 *Amer. J. Sci.* **13**, 223.
Wyckoff, R. W. G. and Muller, J. H. 1927 *Amer. J. Sci.* **13**, 347.

Wyckoff, R. W. G. and Posnjak, E. 1921 *J. Amer. Chem. Soc.* **43**, 2292.
Yost, D. M. and Sherborne, J. E. 1934 *J. Chem. Phys.* **2**, 125.
Yü, S. H. and Beevers, C. A. 1936 *Z. Kristallogr.* **95**, 426.
Zachariasen, W. H. 1926*a* *Z. Phys. Chem.* **124**, 277, 436.
— 1926*b* *Norsk. geol. Tidskr.* **9**, 65
— 1927 *Z. Phys. Chem.* **128**, 417.
— 1930*a* *Z. Kristallogr.* **73**, 1.
— 1930*b* *Z. Kristallogr.* **73**, 141.
— 1931 *J. Amer. Chem. Soc* **53**, 2123.
— 1934*a* *Z. Kristallogr.* **88**, 150.
— 1934*b* *Z. Kristallogr.* **89**, 529.
— 1936 *J Chem. Phys.* **4**, 618.
— 1937 *Z. Kristallogr.* **98**, 266.
Zachariasen, W. H. and Buckley, H. E. 1931 *Phys. Rev.* **37**, 1295.
Zachariasen, W. H. and Mooney, R. C. L. 1934 *Z. Kristallogr.* **88**, 63.
Zachariasen, W. H. and Ziegler, G. E. 1931 *Z. Kristallogr.* **80**, 164
— — 1932 *Z. Kristallogr.* **81**, 92.
Zahn, C. T. and Mills, J. B. 1928 *Phys. Rev.* **32**, 497.
Zambonini, F 1930 *Bull. Soc. Fr. Min.* **53**, 443
Zocher, H. 1920 *Z anorg. Chem.* **112**, 58.

The wave form of atmospherics at night

BY B. F. J. SCHONLAND, F.R.S., J. S. ELDER, D. B. HODGES,
W. E. PHILLIPS, AND J. W. VAN WYK

(*Received* 5 *February* 1940)

[Plates 1-3]

The wave form of all atmospherics received at night from sources within 2000 km. can be accurately described as a ground pulse followed by a series of sky pulses produced by successive reflexions between the ionosphere and the earth, thirty such reflexions being frequently recorded. The time separation between the peaks of these pulses is determined by the distance travelled and the height of the layer. The primary pulse emitted by the source is usually a single complete oscillation of period ranging from 50 to 400 μsec. At distances greater than 500 km. the ground pulse and the first sky pulse merge owing to the shortness of the time interval between them. Differences of amplitude, form and phase between pulses can arise from differences in angle of emission from the parent lightning channel.

The height of the reflecting layer can be determined within ± 1 km. It ranged from 85·5 to 90·5 km. during two winter months, with a mean of 88·0 km.

The distances of the sources as found by analysis of the pulse series were corroborated by independent location with cathode-ray direction-finders.

The reflexion coefficient of the layer for the pulses of longer period exceeded 0·80.

The velocity of the ground pulse where it can be tested is within 0·7 % of that of light.

1. INTRODUCTION

This paper gives an account of an investigation of the wave form of atmospherics received during the night hours and discusses the interpretation of the records obtained.

Two recent studies of atmospherics, largely concerned with wave forms observed by day, have been made by Lutkin in England (1939) and Laby, McNeill, Nicholls and Nickson (1940) in Australia. The Australian workers originally suggested (Laby, Nicholls, Nickson and Webster 1937) that the structure of the "high-frequency" portion of an atmospheric, which by day appears as a damped wave train of gradually increasing wave-length, arises from multiple ionospheric reflexions of a single and simple "pulse" of short duration. In their later and fuller communication (Laby *et al.* 1940) they consider that this explanation cannot well be fitted to the majority of daylight forms without the assumption that the first five half-periods arise directly from the parent lightning discharge.

Lutkin (1939) considers that the whole of the daylight wave form arises from oscillations and multiple discharges in the lightning channel.

The effect of a lightning discharge should, however, be to radiate a single oscillation of the short pulse-like form originally suggested by Laby and his collaborators, and the regular variation with distance of the "wave-lengths" observed in the day forms, first established by Watson-Watt, Herd and Lutkin (1937), definitely suggests that all peaks on the wave form subsequent to the first arise from a propagation mechanism.

It is here shown that the forms taken by all atmospherics at night arise from a simple direct pulse followed by a series of ionospheric reflexions, as was suggested by the Australian workers in their first communication (Laby 1937).

2. METHOD OF OBSERVATION

The observations were made in Johannesburg in 1938 and 1939 at the Bernard Price Institute and in Durban at the Natal University College, each of which was provided with equipment for the recording of wave form and for the direction-finding of sources of atmospherics. The cathode-ray direction-finders were employed to find the distance travelled by the atmospherics examined.

(a) *Aerial, amplifying and recording equipment*

This equipment followed closely upon the lines developed by Appleton, Watson-Watt and Herd (1926). A paraphase amplifier, preceded by a single triode stage, was employed in Johannesburg. The amplifying equipment in Durban was of the type described by Webster (1937).

The oscillographs used were vacuum-focused tubes (Cossor Types 3276 and 3278), and records were made upon paper or film carried on revolving drums with an axial traverse of 1/10 mm. per rev. (Lutkin 1937).

The peripheral velocities used were $7\cdot7 \times 10^2$ and $3\cdot3 \times 10^2$ cm./sec., in Johannesburg and Durban respectively. These were maintained very constant during each run by the use of constant-speed motors and stroboscopic disks. Good recording could be carried out at much higher film speeds. The use of the devices discussed under (*b*) and (*c*) allowed us to run the same film a number of times.

(b) *Automatic brilliancy control*

For satisfactory operation of the traversing drum method of recording, steps must be taken to avoid the troublesome fogging of the film produced by the stationary spot on the oscillograph and by the general fluorescence of the screen One method used by us was to reduce both spot brilliancy and the background fluorescence to a low value and to control the brilliancy automatically by the transient atmospheric itself. A circuit for this purpose was suggested by one of us to Messrs A C. Cossor, Ltd., and developed by them as an automatic brilliancy control. In this device the potential of the modulating cylinder of the cathode-ray tube is controlled by the rate of change of the potential difference, E, applied to the deflecting plates, the cylinder receiving an additional positive potential proportional to the numerical value of dE/dt.

(c) *Anticipatory leader triggering*

A second method of avoiding the effect of the stationary spot and at the same time securing extremely high brilliancy during the passage of an atmospheric was developed from the anticipatory switching trigger employed in Johannesburg for the photography of lightning in daylight (Schonland 1938). This trigger is operated by the rapid train of pulses arising from the stepped-leader process and is connected to the separate trigger unit in the oscillograph in such a way as to raise the spot brilliancy from zero to a high value during the early part of the leader disturbance, in anticipation of the main wave form. The triggered brilliancy used can be extremely high, enough to burn the screen badly if maintained for long, and it was arranged to be 'on' for about 1/50 sec., sufficiently long to cover the recording of the main wave form. The advantage of anticipatory triggering is that it ensures good recording of the smallest disturbances at the start of the main wave form. Cloud discharges were similarly triggered and recorded for a considerable fraction of their duration. All the records reproduced in this paper were obtained with this device, which can be operated by atmospherics from storms at a distance exceeding 3000 km. 'Subsequent stroke' wave forms, since they usually have no leader process preceding them, do not operate the trigger. These were recorded with the automatic brilliancy device described under (*b*) above.

(*d*) *Location of sources of atmospherics and automatic marking*

Individual wave forms were marked automatically, in such a way as to identify the storm centre from which they came. The two direction-finders were first operated, with trunk telephone communication between the stations, for a sufficient time to allow of the fixing of the positions of all sources of atmospherics within 2000 miles. The screen of the direction-finder tube in Johannesburg was then fitted with a metal cover in which was an adjustable slit which could be set to a particular bearing. When the spot of the direction-finding tube moved out along this bearing as a result of the arrival of an atmospheric from a particular source, light passed through the slit and fell upon a photoelectric cell, which then switched on an argon-filled gas relay valve. The pulse of current from this valve passed *via* a condenser C through a low resistance R in the aerial circuit of the wave-form equipment and deflected the spot for a time set by the product CR. A suitable adjustment of the circuit placed this marking deflexion immediately after the record of the wave form of the atmospheric which operated the direction-finding equipment. All wave-form records followed by the characteristic mark were then known to have arisen from atmospherics with the direction of arrival originally selected, and hence from a known storm centre. After one such source had given a sufficient number of atmospherics, the slit on the direction-finder was rotated to select another and a second run was taken on another film. If the original intersections showed that the Johannesburg bearing ran through more than one storm centre, this bearing was not used.

(*e*) *Recording of wave forms of the same atmospheric at two stations, with position of the parent storm*

A system of marking by hand was employed to identify the records of the same atmospheric at both Johannesburg and Durban. A series of short pulses, from 1 to 10 in number, could be passed through the resistance R mentioned in (*d*) by operating a standard telephone call dial, the brilliancy trigger coming into action at the same instant. The operators of the two wave-form instruments were in telephonic communication, and an atmospheric which triggered the oscillographs at both stations was marked by them with the same number. Two more observers, listening on the same telephone line, watched the compass dials of the direction-finders at the two stations and recorded the directions of arrival corresponding to each marked atmospheric A check-up afterwards gave the position of the storm centre responsible for each marked wave form.

(*f*) *Accuracy of measurement*

The time intervals on the wave forms were measured by means of a micrometer eyepiece to the nearest 0·02 mm. With the speeds employed in the camera the time intervals on good records could thus be measured with maximum errors of ± 6 μsec in Johannesburg and ± 12 μsec. in Durban.

3. GENERAL DISCUSSION OF RESULTS

If the radiation field from a lightning discharge is of a simple character, restricted to one complete cycle of field change, which for brevity will be referred to as a 'pulse', the resultant atmospheric wave form should consist of a ground or direct pulse (G) followed by a series of reflected pulses, $S_1, S_2, \ldots, S_n$ (Laby *et al.* 1940). The first sky pulse, S_1, will have undergone a single reflexion at the ionosphere, and the nth sky pulse, S_n, n reflexions at the ionosphere and $n-1$ at the ground. This series is in fact the characteristic form observed by us at night. Examples of the two main types N_1 and N_2, most of which from their amplitude and from the presence of a leader precursor we associate with discharges to ground, are shown in figures 1 and 2. These are drawn from the original records, examples of which are reproduced in plates 1, 2, figures 3–5. In the N_1 forms all the sky pulses are separated by quiet intervals, but in the N_2 forms they merge into one another in the early part of the series, since the primary disturbance has a duration comparable with the pulse separation.

If t_n is the time elapsing between the arrival of the G pulse and that of the nth sky pulse, S_n,

$$ct_n = \sqrt{4n^2h_n^2 + d^2(1+h_n/R)} - d, \tag{1}$$

where c is the velocity of the pulses, h_n the effective height of reflexion for the *n*th sky pulse, *d* the great-circle distance of the source, and *R* the radius of the earth (Laby *et al.* 1940).

This expression allows for the curvature of the earth and the concentric reflecting layer, using an approximation which is valid up to the greatest

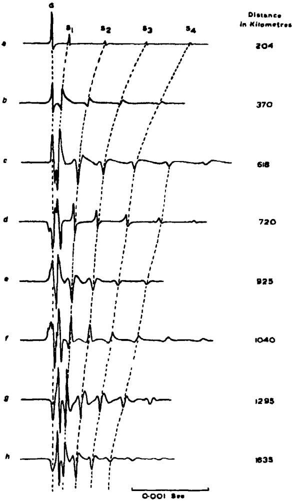


FIGURE 1

distance (3000 km.) employed in our work. It fails, however, to hold for S_1 after a critical distance *d* for which this sky pulse must leave and return to the earth tangentially. This distance is 2130 km. for $h = 90$ km., and 1740 km. for $h = 60$ km. For distances greater than these the first sky pulse can only reach the receiver as a ground pulse travelling along the surface of the earth. Tangential cases for higher orders than the first occur at distances which lie outside our observational range.

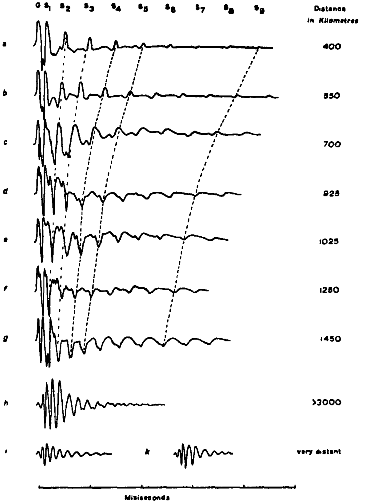


FIGURE 2

It is assumed in equation (1) that c is the same for all the pulses in the series. The simplest method of analysing the data is to suppose initially that this is the case and further that h_n is the same for all values of n.

The set of wave forms shown in figure 6 as tracings and reproduced in plate 2, figure 7 serves as an initial test of these assumptions and of the truth of equation (1). All six wave forms were recorded between 9.30 and 9.32 p.m.

on 19 May 1939. It will be seen from the figure and from table 1 that the values of t_n are practically identical from one atmospheric to the next. Their means are shown in the penultimate line of the table and the calculated values from equation (1) (with h made equal to 88·7 km. and d to 204 km., c being taken as 3×10^{10} cm./sec.) are shown in the last line. The error of measurement is 5 μsec, and the close agreement of observed and calculated values is evidence for the constancy of h and c in this case.

It will be seen from figure 6 that except in the case of atmospheric (f) the ground pulse is the largest of the series. In all cases, as would be expected, the amplitudes of successive sky pulses follow a diminishing series, though this series is not the same for each atmospheric (§ 6).

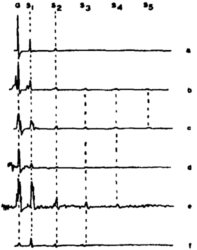


FIGURE 6

Before discussing other tests of equation (1) and the forms illustrated in figures 1 and 2 it is necessary to point out that atmospherics (b) to (e) of figure 1 show a very important feature, which occurs at distances of the order of 400–500 km. This is the merging of G and S_1 into one complex pulse, the early stages of which are illustrated by (a), (b) and (c) of figure 1. This effect, whose cause is discussed in § 10, makes it impossible to use the t values of equation (1) directly, since the G pulse from which they are measured is not clearly distinguishable. We employ instead an expression for the intervals $\tau_1, \tau_2, \ldots, \tau_n$ between successive sky pulses, where

$$\tau_n = t_n - t_{n-1} = \sqrt{[4n^2h^2 + d^2(1+h/R)]} - \sqrt{[4(n-1)^2h^2 + d^2(1+h/R)]}. \quad (2)$$

Some examples of the application of this expression as a test of the theory are shown in table 2. The table includes calculated values based on the same assumptions as regards h and c as before, and the agreement is again extremely good.

The examples in tables 1 and 2 show that with two parameters h and d, the first of which has practically the same value (88·5 km.) in all cases, the reflexion theory accounts satisfactorily for the whole of the wave form of a night atmospheric. The distance d in a number of cases can be confirmed by the use of direction-finders (§ 9). We have examined more than 200 such wave forms, all of which give the pulse series represented by equation (1). The values of h have been accurately found in sixty-six cases and are discussed in § 5.

TABLE 1

(CX 40, figures 6 and 7), $h = 88{\cdot}7$ km., $d = 204$ km.

Atmospheric	Time of arrival after G pulse, in μsec.				
	t_1	t_2	t_3	t_4	t_5
40_1	227	680	—	—	—
40_2	233	691	1223	1783	2355
40_3	232	684	1225	1780	2349
40_4	232	682	1219	1782	—
40_5	219	680	1211	—	—
Means observed	229	683	1222	1782	2352
Calculated intervals	225	687	1221	1782	2355

TABLE 2

Atmospheric	h (km.)	d (km.)		Separation between pulses in μsec.							
				τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	τ_7	τ_8
89, 8	89·0	700	Obs	*	*	313	386	442	484	502	521
			Calc	88	209	313	392	441	480	507	521
90_2, x	89 0	700	Obs.	*	213	309	391	435	478	507	521
			Calc.	88	209	313	392	441	480	507	521
102_p, α	88·5	370	Obs.	140	330	450	505	—	—	—	—
			Calc.	143	338	440	500	—	—	—	—
101, z	88 5	720	Obs.	*	200	305	385	435	460	-	—
			Calc.	89	202	304	375	435	475	—	—
101, y	88 5	925	Obs.	*	*	275	325	385	—	—	—
			Calc.	78	162	255	320	385	—	—	—

* Pulse too broad to measure those intervals accurately.

The last three examples, (h), (j) and (k) of figure 2, are judged from their small amplitudes with high amplifier gain and from the small time intervals between pulses to have travelled distances exceeding 2500 km., a conclusion supported by direction-finding tests made at the time such records were obtained. The detailed discussions of such forms will not be attempted here.

Many of the N_2 wave forms in figure 2 continue the series of reflected pulses beyond the limits of the drawing, which would have to be very large to include the whole series. One such example with reflexions up to the 40th order is sketched in figure 8. The original record, with others of the same type, is reproduced in plate 3, figure 9. It is not at all unusual to find twenty to thirty pulses on these records. The time intervals between pulses become nearly constant towards the end of the series and tend asymptotically in all cases to a value close to 600μsec (§ 4*c*). This interval corresponds to reflexions of such high order that the sky pulses strike the layer at nearly vertical incidence. For example, with $d = 900$ km. the distance between two ground reflexions for the 30th order is 30 km., and the angle of incidence at the ionosphere is 9·5° if $h = 90$ km. The limiting value of the time interval is clearly $2h/c$, and for the above value of 600μsec. these high-order pulses would yield $h = 90$ km., close to the values found in tables 1, 2 and 3 (§ 4) for the low-order reflexions.

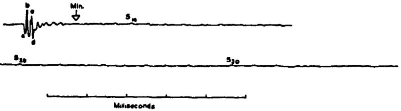


FIGURE 8

This type of wave form was first recorded by Burton and Boardman (1933) in their investigations of atmospheric disturbances on cable and telephone lines. They put forward the same explanation of their origin as has been given here, and with Barkhausen (1930) must be credited with the original discovery of the reflexion mechanism. They showed that such wave forms were responsible for a characteristic audible disturbance on cable lines, known as a 'tweek', whose pitch falls rapidly to a final constant value, 1650 cyc./sec., corresponding to a pulse interval of 605μsec.

4. SEPARATE DETERMINATION OF h AND d

The application of equations (1) or (2) to the large number of wave forms available is tedious and not in itself always sufficiently accurate for a complete test of the reflexion theory. A simpler and more accurate test is to derive the value of h from different parts of the same wave form and to compare the results.

(*a*) *Determination of h*

Owing to the merging of the G and S_1 pulses discussed in § 3, it is convenient to use an expression for h into which the t values for the sky pulses do not enter. If t_p, t_q and t_r are the times of arrival of the pth, qth and rth sky pulses, and if $t_q - t_p = \tau_1$, $t_r - t_q = \tau_2$, it is easily shown from equation (1) that

$$h = \frac{c}{2}\sqrt{\frac{\tau_1 \times \tau_2 \times (\tau_1 + \tau_2)}{(r^2 - q^2)\,\tau_1 - (q^2 - p^2)\,\tau_2}}. \tag{3}$$

This expression does not contain the distance d. If the pulses p, q and r are chosen so as to be separated by large time intervals, the measurements of τ_1 and τ_2 and the value of h derived from them can be arranged to have a small percentage error.

It is assumed in the derivation of (3) that h is the same for all the three pulses making up the triad p, q and r, so that if h varies the value obtained is a mean over the triad. By choosing two sets of triads in the same atmospheric we can test whether h alters as we pass from lower to higher orders. Some tests of this kind are shown in table 3. The limits of accuracy shown are maximum possible errors arising from errors of measurement.

TABLE 3. TEST FOR CONSTANCY OF h BY TRIAD FORMULA

Atmospheric	Distance (km.)	Triad	h (km.)	Triad	h (km.)	Mean h
90, 2x	700	1-4-7	89·5 ± 1·5	4-7-10	87·0 ± 1·2	88·2 ± 1·4
89, 7	635	1-3-5	87·2 ± 3·5	3-5-7	89·9 ± 1·4	88·5 ± 2·4
89, 15	700	1-3-5	93·0 ± 1·1	5-7-9	91·3 ± 1·7	92·2 ± 1·4
104, 1	480	1-3-5	92·0 ± 4·0	5-7-9	91·7 ± 2·0	91·8 ± 3·0
109, 8	700	2-4-7	82·9 ± 4·0	4-7-10	87·4 ± 1·5	85·6 ± 2·7
113, 8	2700	20-30-39	88·8 ± 2·0	26-36-46	88·7 ± 0·7	88·7 ± 1·3
Unweighted means			88·8		89·3	

The results in table 3 indicate that h is sensibly constant for each pair of triads, since the difference in h in each case is within the possible error of this quantity. Further, the unweighted mean of h from the earlier triads (column 4), 88·8 km., is very close to that from the later ones (column 6), 89·3 km.

The constancy of h with increasing order is to be interpreted as a constancy of h with decreasing angle of incidence on the ionospheric reflexion layer and is evidence for specular reflexion of the pulses. For the first order of atmospheric 104, 1 in table 3 this angle is 70°, while for the 46th order of atmospheric 113, 8 it is 18° 30′.

The triad formula frequently yields values of h which have an error of less than 1 km. The separations of the triads shown in the table have been chosen so as to allow of two determinations of h in each case, and thus involve larger errors.

(*b*) *The determination of d*

Having determined h by the above method we can derive the distance d in a number of ways. The most accurate of these again employs the interval, τ, between two widely separated sky pulses, the pth and the qth where $\tau = t_q - tp$. It is easily shown that

$$d^2(1+h/R) = \left[\frac{4h^2(q^2-p^2)-c^2\tau^2}{2c\tau}\right]^2 - 4p^2h^2. \tag{4}$$

As an example we may take the case of atmospheric 89, 7, for which the mean value of h (table 3) is 88·5 ± 2 km For p and q equal to 2 and 4 equation (4) gives $d = 620\,\text{km}$, while for p and q equal to 5 and 7 it gives $d = 640\,\text{km}$.

A simpler method of finding d and of checking the series is to make use of a set of curves giving the values of $\tau = t_{n+1} - t_n$ for the first six terms of equation (2) with various values of h and d. The correct values of these parameters are then obtained by interpolation.

(*c*) *h and d from high-order reflected pulses*

When n is large, $t_n - t_{n-1}$ approaches $2h/c$ asymptotically, according to the relation

$$c(t_n - t_{n-1}) = c\tau = 2h - \frac{d^2(1+h/R)}{4hn(n-1)}. \tag{5}$$

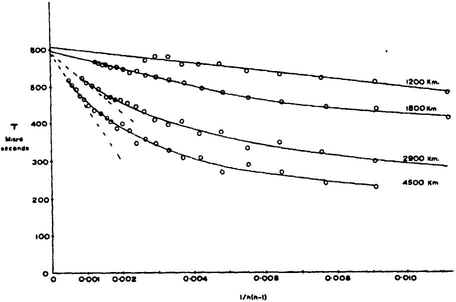


FIGURE 10

If τ is plotted against $1/n(n-1)$ the intercept with the τ axis yields $2h/c$ and hence h, c being taken as the velocity of light. When n is large equation (5) gives a straight line, of slope $\tan^{-1}[-d^2(1+h/R)/4hc]$. Substitution in this expression of the value of h determined from the intercept yields the distance d.

Figure 10 shows four curves in which $t_n - t_{n-1}$ is plotted against $1/n(n-1)$. The intercepts yield values of h varying from 88·5 to 91 km. The variation in slope of the linear parts of the curves with distance is evident. These slopes yield the values of d shown in the figure, which range from 1300 to 4500 km.

Figure 10 shows that the linear part of such a curve is much reduced in extent as d increases. This is because equation (4) depends upon an approximation which for larger values of d requires larger values of n.

5. THE HEIGHT OF THE REFLECTING REGION AT NIGHT

The N_1 and N_2 types of wave form and the cloud-discharge wave forms of type N_1 (§ 8) are not observed by us in the daytime. As noted by Burton and Boardman in the case of 'tweeks', these forms appear at sunset and disappear immediately after sunrise, when they are replaced by distinctly different day wave forms.

When the night wave forms are analysed by the triad formula for the value of the height h of the reflecting layer, it is found that h is remarkably constant. Some examples of values obtained from individual atmospherics recorded during the same runs are given in table 4.

TABLE 4. VALUES OF h FROM DIFFERENT ATMOSPHERICS RECORDED IN THE SAME RUN

Atmospheric	Time and date of run	h (km.)	Mean h (km.)
104, 1	01.00 hr	91·8 ± 2·0	91·6 ± 2·0
104, 2	25. vii. 39	91·5 ± 2·0	
114, 7	05.00 hr	84·4 ± 2·0	
114, 8	27. vii. 39	84·6 ± 2·0	85·6 ± 2·0
114, x		87·0 ± 2·4	
114, 5		86·2 ± 1·2	
116, 5	06.30 hr.	84·9 ± 1·2	
116, x	27. vii. 39	85·3 ± 3·0	85·6 ± 1·0
116, 4		86·5 ± 1·0	

Our night observations extended over a period of two months, from 19 May to 29 July 1939, during the Southern winter. The values of h found during this period are shown in figure 11. All values found during a single run of a few minutes have been combined to give a mean, as in the last column of table 4 above. The data plotted in figure 11 are derived from sixty-six individual atmospherics.

On two nights, as shown by the symbols in the figure, observations were made at regular intervals from sunset to sunrise. The results plotted in the figure refer to observations made both in Johannesburg and Durban, whose values of *h* show no significant difference.

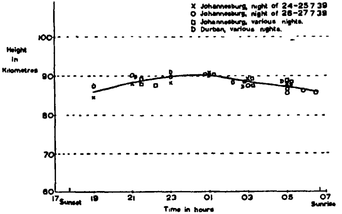


FIGURE 11

From 9 p.m. to 5 a.m. the value of *h* is fairly constant over the whole period, ranging from 85·5 to 90·5 km. with a mean of 88·0 km. There is evidence of a maximum at 1 a.m. with lower values in the early evening and of a fall towards sunrise. Burton and Boardman, from their observations of 'tweeks', have reported values for *h* ranging from 83·5 to 93·2 km. during the hours of complete darkness, together with lower heights at sunset and sunrise. Three night values are given by Laby and his collaborators (1940), 80·5, 81 and 83 km.

6. THE RELATIVE PHASE, FORM AND AMPLITUDE OF THE PULSES IN THE WAVE-FORM SERIES

Though the pulses which make up the wave-form series all arise from the same lightning channel, the following considerations show that they are not necessarily all identical in form or 'phase'.

If the radiation from the original channel is regarded as emitted from a series of linear channel elements, one above the other and each inclined at a different angle to the horizontal, the observed pulse will be the resultant of the summation of the pulses from these elements. The elementary pulses will differ in amplitude and form and in the time of their emission. The contribution of each of them to the resultant pulse will depend upon the angle of emission selected for study. This angle varies according as the resultant pulse is a ground or a sky pulse and is selected by the order of the sky pulse considered.

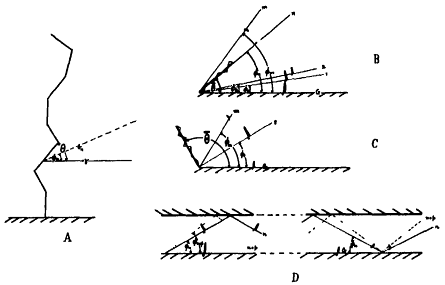


FIGURE 12

Consider an element of the channel which is inclined at an angle θ to the horizontal. Let γ be the amplitude of the elementary ground pulse emitted by it at the angle $\phi = 0$ (figure 12 A), and measured at some small unit distance from it Let σ_n be the corresponding amplitude, measured at the same distance, of its contribution to the nth sky pulse. Then

$$\gamma/\sigma_n = \sin\theta/\sin\overline{\theta-\phi_n}, \qquad (6)$$

where ϕ_n is the value of ϕ for the nth sky pulse. This ratio may pass from positive to negative values as ϕ_n passes through the angle θ, at which point $\sigma_n = 0$. If $\theta > \frac{1}{2}\pi$ the ratio is always positive.

It is clear from equation (6) that when a summation of all such elementary pulses is made we may expect to find considerable variations in the form of the resultant sky pulse with variation in n. As a rule, however, such variations will not be prominent in low-order pulses from fairly distant sources, since the lightning channel is usually nearly vertical and since ϕ_n in such cases is small. Thus in general, as the illustrations in figure 1 show, the low-order pulses are very similar in form to the ground pulse.

With nearer sources or with pulses of high order at any distance we must expect a considerable change in pulse form as ϕ_n increases. An example is shown in figure 1 *b* where the sky pulses are all reversed in phase with respect to the ground pulse This can arise if ϕ_1 is greater than the inclination of the majority of the channel elements to the horizontal

If the actual channel be supposed replaced by a linear channel inclined at an angle $\bar{\theta}$ and $\bar{\theta}$ is $< \frac{1}{2}\pi$, such a reversal of phase will occur for a value $\phi_n \geqq \bar{\theta}$ and the amplitude of the sky-pulse series will pass through a minimum at $\phi = \bar{\theta}$ (figure 12 B) Some examples of this minimum are shown in figures 8 and 9. Another minimum effect occurs early in some N_2 wave forms due to overlapping and interference of long-period pulses. It can, however, be distinguished from this one by the absence of clearly marked separate pulses in low order parts of the series.

If $\bar{\theta}$ is greater than $\frac{1}{2}\pi$ the radiation of the channel as a whole may be greater in the direction of the sky pulses than along the ground and the G pulse amplitude disproportionately small (figure 12 C). An example is shown by atmospheric (*f*) of figures 6 and 7. From table 1 it appears that the G pulse in this case has travelled 204 km and the S_1 pulse 272 km., the angle ϕ_1 thus being close to 45°. If $\bar{\theta}$ had been $\frac{1}{2}\pi$, the reflecting power of the ionosphere 0·80 and the attenuation of the ground pulse negligible, the amplitude ratio G/S_1 should have been 1·80. For the other atmospherics, (*a*) to (*e*) of figure 6, which came from the same source, this is the order of magnitude of the ratio observed but for atmospheric (*f*) it is 0·5. We conclude that $\bar{\theta}$ in this case exceeded 135°. The example illustrates the conclusion that no two atmospherics can give precisely the same amplitude distribution among the reflected pulses of low order. For very high orders, however, there will be little difference between ϕ_{n+1} and ϕ_n and here the decrease in amplitude with order, if no minimum exists, will be set largely by increase of distance travelled and additional loss at reflexion.

7. THE REFLEXION COEFFICIENT OF THE IONOSPHERIC LAYER

If s_{n+p} and s_n are the amplitudes of the $(n+p)$th and the nth resultant sky pulses, produced by the summations of σ_{n+p} and σ_n discussed above, and measured at the same small distance from the parent channel, the ratio of the recorded amplitudes S_{n+p} and S_n will be given by

$$\frac{S_{n+p}}{S_n} = \frac{s_{n+p}}{s_n} \frac{D_n}{D_{n+p}} \frac{\cos\phi_{n+p}}{\cos\phi_n} (r_i \times r_g)^p, \tag{7}$$

where D_n is the distance travelled by S_n, $90 - \phi_n$ its angle of incidence on the layer (figure 12 D) and r_i and r_g the reflexion coefficients of the layer and the ground respectively. It has been pointed out that s_{n+p}/s_n depends on the nature of the parent lightning channel. However, if this channel is nearly vertical and both n and d are large we may expect that s_{n+p}/s_n is not far from unity. Further, under the same conditions, D_n/D_{n+p} and $\cos\phi_{n+p}/\cos\phi_n$ will both be only slightly less than unity. If then we take $r_g = 1$, as is the case for long radio waves, equation (7) can be simplified to

$$r_i \geqq \sqrt[p]{\frac{S_{n+p}}{S_n}}. \tag{8}$$

This equation should give a lower limit to r_i which is not very far from its true value. Some values of r_i derived in this way are given below. They show a reflexion coefficient which lies between 0·73 and 0·93. The values of r_i obtained by measurements on higher frequencies are considerably lower than those given above. Thus Smith-Rose and Barfield (1927) give $r_i = 0{\cdot}11$ for $\lambda = 386$ m. and $r_1 = 0{\cdot}17$ for $\lambda = 479$ m., indicating a rise with wave-length which would be expected on general grounds. The values given in table 5 refer to pulses from N_2 wave forms whose period is of the order of 200 μsec. and which thus have an effective wave-length of 60 km. For pulses of shorter period, which are responsible for the N_1 type of wave form, the reflexion coefficient is smaller and is estimated at 0·50. The smaller number of reflected pulses to be found on N_1 forms is thus a consequence of the shorter period of the primary pulse

TABLE 5. REFLECTING POWER OF THE IONOSPHERE FOR ATMOSPHERIC PULSES

Atmospheric	n	p	r_i (lower limit)
106D	8	10	0·78
	10	4	0·76
	8	4	0·78
84, 3D	8	2	0·73
77, J	15	10	0·93
91, 2J	9	5	0·80

8. WAVE FORMS FROM CLOUD DISCHARGES

The wave forms so far discussed have been those arising from discharges between cloud and ground. A second and more numerous class has been attributed (Schonland, Hodges and Collens 1938; Lutkin 1939) to discharge within the cloud or into the air outside the cloud. Lutkin further divides this 'cloud' class into two groups. The first of these subdivisions (group II) consists of the leader process only and is to be identified with discharges of the stepped leader type which have no return stroke since they do not reach the ground. The second (group III) consists of a set of complete wave forms each similar to the single series obtained from ground discharges but of smaller amplitude and separated by very short quiet intervals. This class, we would suggest, is to be identified with the set of successive dart discharges which has been observed photographically as an alternative to the stepped leader process in the case of cloud-air discharges. Examples of these two classes as observed at night are given in plate 3, figure 13.

As would be expected, both groups II and III exhibit the series of reflected pulses which is characteristic of ground discharge forms (group I). The existence of the series is often difficult to establish in the case of the stepped leader form, since each leader ground pulse is superimposed upon the sky pulse series of its predecessor. Where, however, a prominent pulse occurs, its series can always be followed for some distance along the record. In the group III cloud wave forms the individual series are separated and correspond exactly to the ground-discharge forms of larger amplitude. On many occasions we have been able to establish by direction-finding intersections that wave forms of group III have originated from storms which have given group I records on the same film. Comparison of the time separations of the pulses in the two groups then shows no difference between them.

The duration of the pulses found on group III forms at night is very short, as would be expected if they originated in short dart streamers. The wave form is therefore of type N_1 with clearly separated reflected pulses whose amplitude dies down more rapidly than in the usual N_2 form. As a consequence the series does not show very many reflexions. Lutkin (1939) has reported a similar effect by day, the number of 'oscillations' making up his group III wave forms being less than the number observed on group I forms. Since the reflexion mechanism would at first sight be expected to give rise to the same number of 'oscillations' whatever the nature of the discharge, he concludes that this difference offers evidence against the reflexion theory and in favour of the view that the whole wave form is produced at the parent discharge itself. It has, however, been shown that the number of reflected pulses which can be observed is dependent upon the duration of the primary pulse, and if this is short in group III forms by day as it is by night the reflexion coefficient will be less than usual and the series will terminate earlier than usual.

9. WAVE FORMS OF THE SAME ATMOSPHERIC AT TWO DIFFERENT DISTANCES

The arrangements used to obtain wave forms of the same atmospheric in both Johannesburg and Durban have been described in § 2 (*e*). An example is shown in figure 1, where (*g*) and (*h*) come from the same lightning flash, (*g*) having travelled 1295 km. to Johannesburg and (*h*) 1635 km. to Durban. These distances were roughly confirmed by direction-finding observations. Many other examples of the same kind establish the fact that the temporal structure of the wave form of the atmospheric is a function of the distance travelled by it. When the wave form at one station is of the very distant type (§ 3) shown in figure 2 (*h*, *i*, and *k*) it is found that a similar form is recorded at the other station, the time separation between pulses being practically identical in the two wave forms. In such cases the direction-finders give bearings which indicate that the source is at practically the same distance from both stations

It is not usually possible to rely upon frame aerial direction-finders for the location of sources of atomspherics at night, owing to well-known polarization errors. Occasionally, however, successive direction-finder intersections were so consistent as to indicate that night errors were small or absent. In such cases the great circle distances obtained from the direction-finders were in satisfactory agreement with those derived from an analysis of the wave forms themselves. Some examples are given in table 6 below.

TABLE 6

Atmospheric	Distance from Johannesburg km.		Distance from Durban km.	
	By D.F.	By wave form	By D.F.	By wave form
94, 2, 1	720	800	360	400
94, 3, 4	770	965	455	520
94, 3, 8	1400	1800	*	*
94, 3, 1	*	*	455	570
89, 7	650	660	*	*
114, 1	1500	1300	*	*
114, 2	1250	1300	*	*
114, 3	1100	1020	*	*

* No wave-form record made.

As would have been expected from the analyses of wave form which have been already discussed, the results of this section support the conclusion that the wave form of atmospherics observed at night is due to propagation conditions and is not imposed at source.

10. THE VELOCITIES OF GROUND AND SKY PULSES

The agreement between the observed pulse separations and those calculated on the supposition that the velocity c is the same for all sky pulses is so good as to make it clear that no real variation in c can be established on the present evidence. It must therefore be taken for all sky pulses to be close to that of light. The velocity of the G pulse is less easy to establish since this pulse is difficult to identify after it has travelled 500 km. when it merges with S_1 and is considerably attenuated. Table 1 deals with cases where G and S_1 were clearly separated and G travelled 204 km. The agreement between the observed and calculated value of the $S_1 - G$ interval is such that we conclude that the time taken by G, 680 μsec., is that which it would have taken if its velocity were the same as that of S_1 to an accuracy of 5 μsec. Thus in this case the velocity of the G pulse was within 0·7 % of that of light. This conclusion is identical with that previously reached on similar evidence by Laby (1940).

This result does not preclude the possibility of a larger decrease in the velocity of the G pulse over greater distances. Though we suspect this, we have at present no trustworthy evidence for it. The observed merging of G and S_1 may be entirely due to the shortness of the time interval between their arrival (115 and 77 μsec. respectively for distances of 500 and 1000 km.) compared with the duration of the pulses themselves (50–400 μsec.).

This merging of G and S_1 makes it impossible to obtain any data on the attenuation of the G pulse with distance.

11. THE SHAPE OF THE PULSES IN THE WAVE-FORM SERIES

Figure 14 shows the three types of pulse observed in the wave-form series. Type 1 is commonly found in N_1 forms and has a total duration of 50–100 μsec. Type 2 is longer in duration, ranging from 100 to 250 μsec., while type 3 is still longer, up to 400 μsec. 1 and 2 are found in N_2 forms and show reflected pulses of high order.

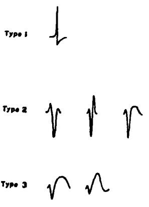


FIGURE 14

The details of the origin of these forms will not be discussed here, but it may be pointed out that their durations correspond closely to the duration of the main return stroke portion of the lightning discharge. Our records rarely show cases of pulses of more complicated type or of longer duration than those stated. They therefore give no support to the view (Lutkin 1939) that subsequent discharge components in the return stroke channel frequently produce appreciable radiation fields. In the few cases where the G pulse can be observed separately from S_1 it is found to be of type 1. All the examples in figure 6 are of this type. It is thus possible that types 2 and 3 arise either from the additional contribution which is made in the case of sky pulses by horizontal parts of the discharge channel or from distortion effects on reflexion.

12. Discussion

In the course of this work many hundreds of night atmospheric wave forms have been examined. All these show the clearest evidence that their pulse series arises from repeated reflexions between the ionosphere and the ground. With very few exceptions the pulses involved are of the simple form and short duration discussed in § 11, indicating that the primary disturbance arose from a simple rise and fall of current in the lightning channel without subsequent oscillations.

The only really certain exception to this rule which we have found is the atmospheric sketched in figure 8 and illustrated in plate 3, figure 9*b*, and noteworthy for having at least forty reflected pulses in its series. The order of any particular pulse in this series can be obtained independently by a simple adaptation of the equations discussed in this paper, in which n as well as h and d are treated as unknowns, and some of these orders are marked in figure 8. It is then found that the S_1 pulse of the series (G having been much attenuated by its passage over 1800 km. and in any case having merged with S_1) is that marked *cd*. The pulse *ab* which precedes it is an S_1 pulse from a different source which does not give rise to any high order reflexions. Full consideration of such exceptional cases must be reserved until a sufficient number have accumulated but it may be pointed out that the absence of high order reflexions from *ab* and the very large number of such reflexions from *cd* would both be explained if *ab* arose from a vertical channel and *cd* from a nearly horizontal one. From this one can suggest that all wave forms showing high-order reflexions arise from lightning channels with a considerable horizontal portion such as is often formed as the return stroke enters the cloud-base. This suggestion would link up with the fact that the primary pulse is of longer duration than usual.

While the majority of our G pulses are very simple in form, a few show a rather intricate fine structure. Two examples are illustrated in plate 3, figure 15, from which it will be seen that the fine structure is gradually lost

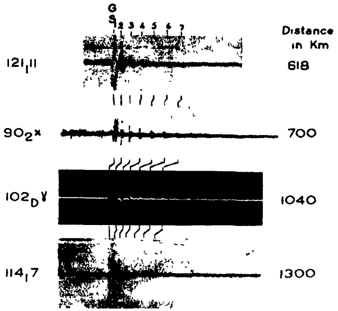


FIGURE 3

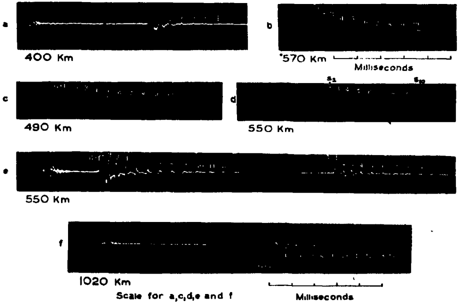


FIGURE 4

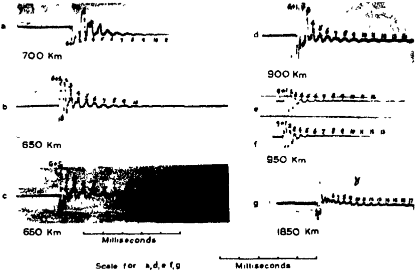


FIGURE 5

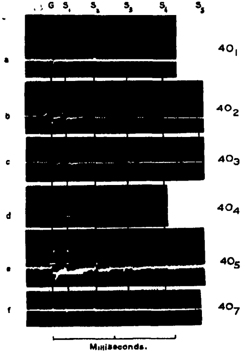


FIGURE 7

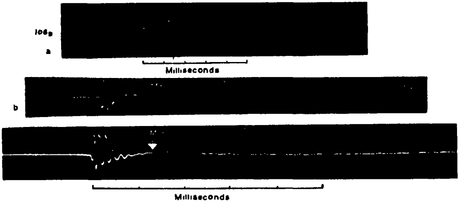


FIGURE 9

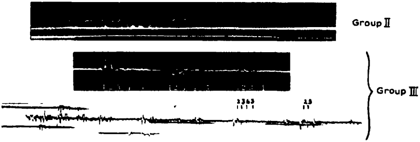


FIGURE 13

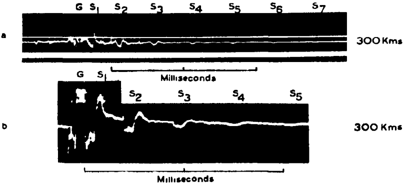


FIGURE 15

in subsequent reflexions, as would be expected from its short period and consequent low coefficient of reflexion.

The general similarity of wave forms from the same source has already been noted by workers in the case of daylight atmospherics and follows as a natural consequence of the reflexion mechanism. Atmospherics *b* and *c* and atmospherics *e* and *f* of plate 2, figure 5 illustrate this similarity very clearly, minute details being reproduced in each pair with remarkable fidelity, though in each case the two wave forms were recorded about a minute apart. (The time scale for *b* is uncertain since it was taken at the close of a run.) Such similarity in detail requires of course a certain general similarity in the shape of the primary pulse and consequently in the slope and duration of the parent discharge. It is, however, impossible to ascribe it to a close similarity in the 'Malan' components of the two discharges (Malan and Collens 1937) and the phenomenon is only explicable on the basis of the reflexion mechanism.

The results discussed in this paper were obtained during the winter months. More recent work indicates that similar night wave forms can be obtained in summer.

13. *Application to atmospheric wave forms observed by day*

The transition from night to day will alter the nature of the wave form of an atmospheric in three ways. In the first place, since the height of the reflecting layer is lower by day, the time separation of the pulses at a given distance will decrease. Secondly the effective reflexion coefficient will be less, owing to greater absorption, and fewer reflected pulses will be observed in the series. Finally, since the conditions may be no longer approximate to those of specular reflexion, the reflected pulses themselves may lose their sharp character and be changed in form and duration. As a consequence the series, unless observed close to the original source, may simulate a damped train of waves.

The first and last of these differences will be of vital importance to the proper interpretation of the wave form, since the G and S_1 pulses will merge by day at a distance which is much smaller than 500 km. Equation (1) for the reflexion series $G, S_1, S_2, \ldots, S_n$ will then hold for atmospherics from comparatively near storms with a correct value of h, but since the series owing to merging and to attenuation of the G pulse will become $S_1, S_2, \ldots, S_n$ further away, calling it G, S_1, S_2 will give an incorrect and larger value of h. Before our night observations were analysed it was thought from our day observations that this apparent increase in the day value of h with distance was real and it was suggested that it arose from a reduction of h in the neighbourhood of the parent thunderstorm (Schonland, Elder, van Wyk and Cruickshank 1939). The discovery of the merging of G and S_1 at night makes this hypothesis unnecessary, and our day observations when analysed anew give a value for h of about 60 km. irrespective of the distance travelled by the atmospheric.

Our thanks are due to the Postmaster-General of the Union of South Africa and to Colonel F. Collins, Under-Secretary for Telegraphs, for generously providing telephone facilities for this work.

REFERENCES

Appleton, E. V., Watson-Watt, R A. and Herd, J. F. 1926 *Proc Roy. Soc.* A, **111** 615

Barkhausen, H. 1930 *Proc Inst. Radio Engrs, N.Y*, **18**, 1155

Burton, E and Boardman, E 1933 *Proc Inst. Radio Engrs, N.Y*, **21**, 1476.

Laby, T. H., McNeill, J. J., Nicholls, F. G. and Nickson, A F B. 1940 *Proc. Roy. Soc.* A, **174**, 145.

Laby, T. H., Nicholls, F. G, Nickson, A F and Webster, H. C. 1937 *Nature, Lond.*, **139**, 837.

Lutkin, F. 1937 *J Sci. Instrum.* **14**, 209.

—— 1939 *Proc. Roy Soc.* A, **171**, 285.

Malan, D. J and Collens, H. 1937 *Proc. Roy Soc.* A, **162**, 175

Schonland, B F J 1938 *Nature, Lond*, **141**, 115.

Schonland, B. F. J, Elder, J. S., van Wyk, J and Cruickshank, G 1939 *Nature, Lond*, **143**, 893.

Schonland, B. F. J, Hodges, D. B and Collens, H. 1938 *Proc Roy Soc* A, **166**, 56

Smith-Rose, R L and Barfield, R H. 1927 *Proc Roy Soc* A, **116**, 682

Watson-Watt, R. A, Herd, F. F. and Lutkin, F 1937 *Proc. Roy. Soc.* A, **162**, 267.

Webster, H. C. 1937 *Proc. Phys. Soc.* **49**, 654.

A theoretical study of a possible model of paramagnetic alums at low temperatures

By J. A. Sauer and H. N. V. Temperley

(*Communicated by R. H. Fowler, F.R.S.—Received* 29 *September* 1939.—*Revised* 23 *April* 1940.)

An attempt is made to examine theoretically the properties of paramagnetic alums at low temperatures. The model taken is a lattice of freely suspended magnets, all interactions except purely magnetic being neglected. Even with this simplification it is impossible at present to make rigorous calculations of the partition function, either on classical or quantum lines. A simple model is proposed, which is really a generalization of the Bragg-Williams theory enabling one to take account of the effect of a magnetic field. The few configurations whose energies are known are used to fix arbitrary constants in the expression assumed for the energy. The theory predicts that the state of lowest energy is either a spontaneously magnetized state for a long thin specimen, or a state in which alternate *rows* of magnets point in opposite directions for a sphere, spontaneous magnetization appearing in an ellipsoid with an eccentricity greater than a certain critical value. The transition curve bounding the region in which the antiparallel state is stable consists partly of a line of Curie points corresponding to transitions of the second order, passing smoothly into a line of critical points corresponding to a transition of the first order. The effect of shape on the magnetic properties of the specimen seems to be experimentally verified, but the rough nature of the theory prevents it being more than qualitative.

Introduction

A paramagnetic alum consists of a crystal containing a large number of molecules of water of crystallization. The paramagnetic ions are therefore some distance apart and their interactions are therefore small. The structure of the crystal is usually either face-centred or body-centred cubic, so that we may expect the crystalline field to have a high degree of symmetry. It therefore seems permissible as a first approximation to regard a paramagnetic alum as a cubic lattice of freely suspended magnetic dipoles. It is precisely this fact that the ionic magnets are so nearly free that makes the paramagnetic alums so valuable for obtaining very low temperatures. In what follows we shall neglect all interactions except the purely dipole-dipole interactions between the paramagnetic ions themselves. Even with this simplification, the problem remains exceedingly complicated, owing to the peculiar form of the interaction between the dipoles, the interaction energy being strongly dependent on the inclination of the dipoles to the line joining them, as well as on their inclination to one another. A well known consequence of this is that, if we have two dipoles pointing in the same direction along the line joining them, they are in a state of minimum potential energy, and there is a potential barrier even between this state and the state in which both dipoles are reversed. Temperley (1940) has shown that this peculiarity may perhaps explain the curious hysteresis effects discovered in iron alum at temperatures above the Curie point by Shire and Barkla (1939).

It is generally accepted that ferromagnetism in metals is due to exchange interactions. Unlike the dipole forces, the interaction between spins coupled only by exchange forces depends only on their inclination to one another. Another difference is that exchange energies fall off practically exponentially with distance, whereas dipole-dipole interactions only fall off as the inverse cube of the distance. Thus, in a metal the exchange interactions are very large compared with the purely magnetic interactions, but in a paramagnetic alum this situation is probably reversed With exchange interactions only, it is a good approximation to neglect all interactions except those between nearest neighbours, whereas if we orient magnetic dipoles properly, we can get a state in which the magnetic field at a given point diverges logarithmically with an infinite crystal, which means that size and shape may be expected to be important. Owing to these fundamental differences, we shall not be surprised if the properties of a paramagnetic salt below its Curie point are not the same as those of a ferromagnetic metal.

THE CONFIGURATIONAL ENERGY

In spite of the fact that the form of the interaction we have assumed is perfectly definite, it has so far proved impossible, owing to the difficulties we have already mentioned, either to enumerate states of equal energy, or even to calculate the energies of specified states, though Sauer (1940) has obtained expressions for the energies of certain special configurations. From this it is obvious that anything approaching a rigorous statistical treatment is, as usual, impossible, and it is necessary to make fairly drastic assumptions. Sauer's (1940) results suggest that the state of lowest energy for a spherical specimen is one in which alternate rows of dipoles point in opposite directions. (We are assuming for simplicity that each spin has only two possible eigenstates.) We also assume that the energy of a given state is, to a first approximation, a function of the parameters S_1, S_2 defined thus: We divide the spins into two sets according to their direction in the lowest energy state and let S_1, S_2 be the number of spins in the two sets which have to be reversed in order to arrive at the state (S_1, S_2), starting from the antiparallel state of lowest energy. Now, if S_1 is different from S_2 the specimen has a net magnetization in one direction, and we get a term in the energy $-2\mu H(S_1 - S_2)$ due to the external magnetic field H, and a further term due to the appearance of magnetic poles on the specimen. If we assume the Lorentz form for the 'local field', an assumption that we shall discuss later on, we get a further term in the expression for the energy of the form $-2\mu^2\nu(S_1 - S_2)^2$, where ν is given by $\frac{4\pi}{3} - n$ and n is the demagnetizing coefficient along the longest axis of the specimen. Next, consider a state in which $S_1 = S_2$. The two terms we have just found will vanish and the departure of the specimen from the lowest energy state is measured by $(S_1 + S_2)$. The simplest assumption we can make about the change of energy is the Bragg-Williams approximation. According to this, the rate of change of the energy with $(S_1 + S_2)$ will be proportional to $(S_1 + S_2)$ for $(S_1 + S_2)$ small, but will fall off as $(S_1 + S_2)$ becomes larger. If N is the total number of spins, and we have $S_1 = S_2 = \frac{1}{2}N$ it is clear that we recover the original antiparallel state. The simplest form that we can assume for the energy change that shall be consistent for both small and large $(S_1 + S_2)$ is a term proportional to $(\frac{1}{2}N - \{S_1 + S_2\})^2$. It is not difficult to see that this assumption is precisely equivalent to the Bragg-Williams approximation, the quantity $(\frac{1}{2}N - \{S_1 + S_2\})$ corresponding to their long-range order.

So far, the introduction of a Bragg-Williams type of approximation is exceedingly tentative. We have only made our formula check up for very special values of S_1 and S_2. Let us put $S_1 = S_2 = \frac{1}{4}N$. This state corresponds to a *disordered* state in the ordinary theory, and it will be such that the original antiparallel structure is completely broken up.

Since $S_1 = S_2$ this state has no net magnetization, and it follows that there is also no trace of the spontaneously magnetized state. Since both types of ordered structure are broken up, we should expect that this state has no interaction energy. If we consider the spontaneously magnetized state for the case of, e.g., a spherical specimen, we find that the interaction energy is also zero. (This result can be established by writing down the expression for the total interaction energy and applying symmetry considerations, and it also follows if we assume the Lorentz expression for the local field, but it is emphasized that no special assumption about the local field need be made.) If we inspect the expression for the energy of a spherical specimen that we have assumed, we find indeed that it predicts equal energies for the parallel and disordered states. Thus, we have succeeded in checking our formula in the middle of the ranges of S_1 and S_2 as well as at the ends. Sauer's (1940) calculations give the value $1{\cdot}8N^2\mu^2$ for the energy required to change the antiparallel state into the parallel state in a spherical specimen. This value holds for either a body-centred or face-centred cubic lattice. If we take this value, our expression for the configurational energy in terms of S_1 and S_2 is

$$E(S_1, S_2) = -7{\cdot}2\mu^2(\tfrac{1}{2}N - \{S_1+S_2\})^2 - 2\mu H(S_1-S_2) - 2\mu^2\nu(S_1-S_2)^2, \quad (1)$$

where N is the number of ions per c.c. and μ their magnetic moment.

THE PARTITION FUNCTION

Since S_1, S_2 each run from zero to $\frac{1}{2}N$, it follows that the statistical weight of the state (S_1, S_2) is

$$\varpi(S_1, S_2) = \frac{\frac{1}{2}N\,!}{(\frac{1}{2}N-S_1)!\,S_1!}\,\frac{\frac{1}{2}N\,!}{(\frac{1}{2}N-S_2)!\,S_2!}. \quad (2)$$

The partition function is given by

$$f(S_1, S_2) = \sum_{S_1S_2} \varpi(S_1, S_2)\exp\left(-E(S_1, S_2)/kT\right). \quad (3)$$

We now apply the usual procedure of searching for the maximum term with regard to S_1, S_2, since this term corresponds to the equilibrium state.

We get two implicit equations for S_1 and S_2 in terms of H, T and ν

$$S_1 = \tfrac{1}{2}N\,\frac{1}{1+\exp\left[\{7{\cdot}2N\mu^2 - 14{\cdot}4\mu^2(S_1+S_2) - 2\mu H - 4\mu^2\nu(S_1-S_2)\}/kT\right]}, \quad (4)$$

$$S_2 = \tfrac{1}{2}N\,\frac{1}{1+\exp\left[\{7{\cdot}2N\mu^2 - 14{\cdot}4\mu^2(S_1+S_2) + 2\mu H + 4\mu^2\nu(S_1-S_2)\}/kT\right]}. \quad (5)$$

By combining these two equations we get

$$x = N - 2(S_1+S_2) = N\,\frac{\sinh(\lambda x/kT)}{\cosh(\lambda x/kT)+\cosh(2\mu F/kT)}, \quad (6)$$

$$I = 2\mu(S_1-S_2) = N\mu\,\frac{\sinh(2\mu F/kT)}{\cosh(\lambda x/kT)+\cosh(2\mu F/kT)}, \quad (7)$$

where $\lambda = 7{\cdot}2\mu^2$ and $F = H + \nu I$. If H and ν are both zero, equation (6) reduces to the exact form of the Bragg-Williams theory for the case in which two components of an AB alloy are present in equal proportions. We now study the more general case in which H and ν differ from zero.

Equation (6) is satisfied automatically if $x = 0$. If we assume for the moment that the configuration corresponding to this case is stable, equation (7) reduces to

$$I = N\mu \tanh(\mu F/kT). \tag{8}$$

Thus, the assembly behaves according to the ordinary Weiss theory of ferromagnetism, provided that $x = 0$, i.e. that the antiparallel state is not stable. For high temperatures we have a *disordered* state, as in the ordinary Bragg-Williams theory, and for large fields a state in which the specimen is magnetized to saturation. These results are what we should expect physically, and we deduce that the antiparallel state can only be stable, if at all, if both H and T are small. We therefore have three distinct states, and wish to determine the phase diagram. As we have three independent variables H, T and ν, it would seem that we require a three-dimensional phase diagram.

Fortunately, matters can be simplified if we consider the two variables T and F, where $F = H + \nu I$. If we do this, we can restrict our phase diagram to two dimensions without suppressing any information * This arises from the fact that equations (6) and (7) only involve H and ν through F It may also be pointed out that this treatment removes the necessity for making assumptions about the local field, since we might also develop the theory by considering that each dipole was subject to the total field F. It will be necessary to reintroduce some assumption about F, since we cannot measure it, but can only measure H and I. F is a more fundamental physical quantity than H, but equations involving F cannot be interpreted experimentally.

THE CRITICAL TEMPERATURE

If we consider F and T as the independent variables, the behaviour of the assembly is governed by equations (6) and (7) and for zero F we get a critical temperature corresponding to a phase change of the second order, given by $T_c^\circ = \frac{1}{2}N\lambda/k$. For F small, the equation determining the critical temperature is obtained by differentiating equation (6) with respect to x and then putting $x = 0$ which gives us

$$T_c = \frac{N\lambda}{k} \frac{1}{1 + \cosh(2\mu F/kT_c)}, \tag{9}$$

which can be solved numerically. There is no re-entrant portion in a graph of the right-hand side of equation (6) as long as $\cosh(2\mu F/kT_c) \leqslant 2$, but a re-entrant portion appears for larger values of F. The critical temperature will then no longer be given by equation (9), but is determined instead by the rule of equal areas. In other words, we now have a first-order transition involving a finite latent heat, and discontinuous jumps in S_1 and S_2. We can continue the critical curve graphically by using the usual rule of equal areas. This portion of the critical curve rapidly approaches the vertical line given by $4\mu F = N\lambda$.

---

* We are very grateful to Mr A. H. Wilson for pointing this out.

We have now succeeded in finding the critical curve within which the antiparallel state can exist. Equation (7) shows that outside the antiparallel region I or $(S_1 - S_2)$ is a perfectly smooth function of both F and T, while it is fairly easy to see that this must also be true inside the antiparallel region, for, since there is no discontinuity in $(S_1 + S_2)$ there can be no discontinuity in $(S_1 - S_2)$ either. There is therefore no sharp critical curve separating the disordered state from the spontaneously magnetized state, but the transition is gradual and continuous. The whole critical curve is plotted in diagram 1 together with two sets of contours along which either x or I is constant. These can easily be computed from equations (6) and (7).

THE EFFECT OF VARYING THE SHAPE OF THE SPECIMEN

If we could observe F experimentally, the shape of the specimen would not be relevant, since any changes in ν are automatically taken care of. In practice, we observe H and I. If we assume the Lorentz local field, we have $F = H$ for a sphere, and figure 1 will show the critical curve If we consider only the case in which H is directed along the longest axis of the specimen (ν positive), any other directions of H being outside the scope of this theory, it is evident that, as ν increases, the term $-2\mu^2\nu(S_1 - S_2)^2$ in the energy favours the *parallel* state at the expense of the other two states. We shall therefore expect the region on the $(H - T)$ diagram within which the antiparallel state can exist to shrink as ν increases. It is not easy to calculate the exact form of the critical curve for finite ν, but we can deduce that, on account of the form of the energy, the field H_c° which makes the parallel and antiparallel states equally stable at absolute zero is given by $\frac{1}{2}N\mu(3{\cdot}6 - \nu)$. This vanishes for $\nu = 3{\cdot}6$ corresponding to an ellipsoid of axis ratio approximately 6·1. The antiparallel state cannot exist at all for a more elongated specimen than this. The critical temperature for the appearance of spontaneous magnetization is then given by equation (8) and is equal to $N\mu^2\nu/k$. For H zero and ν less than 3·6 the critical temperature is independent of ν. This follows from the fact that for H zero, the critical temperature corresponds to a transition between the antiparallel and disordered states, for both of which we have $(S_1 - S_2)$ zero from symmetry considerations if H is zero. The term containing ν is then irrelevant.

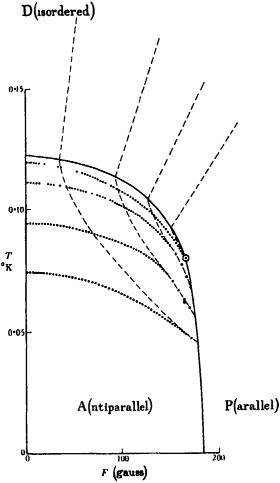


FIGURE 1

○ No latent heat on the transition curve above this temperature

... .. Curves of constant H. - — -- - Curves of constant I

AN APPROXIMATE TREATMENT OF THE EFFECT OF CHANGE OF SHAPE

We can give a slightly different, and perhaps clearer picture of the effect of varying the shape of the specimen, by assuming that for *any* values of H and T, the partition function for, e.g., the antiparallel state may be approximated to by putting S_1 and S_2 exactly equal to zero. We make analogous assumptions for the other states. This treatment is of course inaccurate at any finite temperature. If we require the approximate equilibrium curve between two states, we equate the idealized partition functions. We have as usual

$$\log f(S_1, S_2) = -E(S_1, S_2)/kT + \log \varpi(S_1, S_2).$$

We may tabulate the three states as follows.

State	Energy	$\log \omega(S_1 S_2)$	$\log f(S_1 S_2)$
A	$-1{\cdot}8N^2\mu^2$	0	$1{\cdot}8N^2\mu^2/kT$
D	0	$N \log 2$	$N \log 2$
Q	$-N\mu H - \frac{1}{2}N\mu^2\nu$	0	$(N\mu H + \frac{1}{2}N\mu^2\nu)/kT$

For A we have $S_1 = S_2 = 0$ or $S_1 = S_2 = \frac{1}{2}N$.

For D we have $S_1 = S_2 = \frac{1}{4}N$

For P we have $S_1 = \frac{1}{2}N S_2 = 0$, or $S_1 = 0$, $S_2 = \frac{1}{2}N$.

Inspection of these partition functions gives us the following equilibrium conditions:

Between A and D $\quad kT \log 2 = 1{\cdot}8N\mu^2$. (10)

Between A and P. $\quad H + \frac{1}{2}N\mu\nu = 1{\cdot}8N\mu$. (11)

Between D and P $\quad kT \log 2 = \mu H + \frac{1}{2}N\mu^2\nu$. (12)

These three straight lines are plotted in figure 2 for the cases $\nu = 0$ and $\nu = 1{\cdot}8$ and the more accurate critical curve for $\nu = 0$ is also plotted. It will be seen that the approximate critical curve for $\nu = 0$ is a fair representation of the more accurate curve. The feature we have already predicted, that the critical temperature for H zero is independent of ν for values less than 3·6 appears here also. It is reasonable to suppose that the approximate treatment is also fairly accurate for non-zero ν. The appearance of a definite boundary between the phases P and D, which the accurate treatment shows to be spurious, is clearly due to the fact that we have idealized the phases excessively Everywhere except in the antiparallel region the behaviour of the specimen is governed by equation (8), and, in spite of the fact that this is an implicit equation for I, it is well known that I is a continuous function of H and T, except possibly for H zero and $T < N\mu^2\nu/k$ Owing to the existence of the antiparallel state, we see that we can only get a sharp boundary between P and D if we have H zero, $\nu > 3{\cdot}6$ and $T < N\mu^2\nu/k$.

JUSTIFICATION OF THE ASSUMPTION OF THE LORENTZ LOCAL FIELD

An obvious criticism is that these theoretical critical temperatures are very high. If we consider iron alum, we have μ equal to 5 Bohr magnetons and $N \sim 2{\cdot}2 \times 10^{21}$ ions per c.c. This gives us a critical temperature of 0·14° K, compared with the observed value of 0·034° K (Kurti, Laine and Simon 1937). We have neglected two points in our simplified treatment. We have assumed the full value of μ while still allowing only two eigenstates for the ions, which means that we have a hyperbolic tangent in (8) instead of the proper Brillouin function. For a magneton number of 5, the correction to the critical temperature is a factor of 7/15. If we apply this

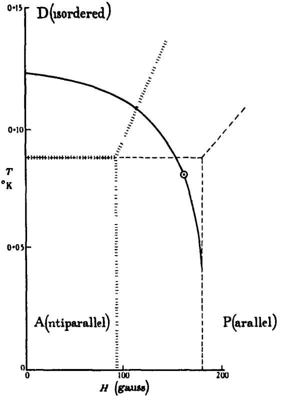


FIGURE 2

——— Accurate transition curve for $\nu = 0$.
- - - - - Approximate transition curve for $\nu = 0$.
ıııııııııı Approximate transition curve for $\nu = 1{\cdot}8$

here, we reduce the theoretical critical temperature to about twice the observed value. This further discrepancy was for a time unexplained, and has been used as an argument against the validity of the Lorentz local field.

The discrepancy was removed by Debye (1938) by taking account of the Stark splitting of the ionic levels due to the crystalline field. When no magnetic field is present, the sixfold degenerate level is split up into a fourfold and a twofold level, with a separation of the order $0{\cdot}06k$. If a magnetic field is applied as well, the degeneracy is removed, but it is possible to obtain approximate expressions for the energies of all six levels. We can thus write down the partition function, and obtain the magnetic moment by differentiating with regard to H. If now we insert the Lorenz local field $\frac{4}{3}\pi I$ in place of H, we obtain an implicit equation for I in zero field. If we study this equation numerically, we conclude that non-zero values of I are possible if T is less than the critical value 0·038° K., in satisfactory agreement with the observed value.

THE EXPERIMENTAL EVIDENCE

There seems to be sufficient, though scanty, evidence for the existence of spontaneous magnetization below the Curie point (Kurti, Laine and Simon 1937, Ashmead unpublished). Simon (unpublished) reports also that the Curie temperature does not vary much with shape, while Ashmead reports that there is no spontaneous magnetization in a sphere, except for a small effect possibly due to impurities etc , but that spontaneous magnetization *does* appear in an ellipsoid with (for iron alum) axes in the ratio of 1·8 to 1. This only agrees qualitatively with the predicted value of 6 to 1, but part of the discrepancy is perhaps due to the fact that we have assumed that each ion has only two states The theory seems to be in rough agreement with the known facts, but there has been very little work on temperatures below the Curie point. Nor indeed is it easy to see what further predictions we could make from the theory, though it might perhaps be worth looking for the change from a second order to a first order transition as we go along the critical curve A jump in $(\partial I/\partial H)$ should change suddenly to a jump in I above a value of H of the order of 150 gauss.

DISCUSSION

Van Vleck's treatment (1937) predicts that dipole-dipole interactions are incapable of producing ferromagnetism. Apart from the experimental evidence, it must be pointed out that this treatment is based on an expansion of the partition function in powers of $1/T$, which must break down at sufficiently low temperatures. We arrive at a similar conclusion if we use the Onsager local field instead of the Lorentz, but there are theoretical objections to both forms Apart from the famous assumption (common to both, which we shall not attempt to discuss) that it is legitimate to scoop out a small spherical cavity in the specimen and fix one's attention on a single dipole in the centre, there are other difficulties Onsager's theory attempts to take account of two features neglected by Lorentz. In the first place, the scooping out of the cavity would affect the polarization of the material, and so would the orientation of the single dipole at the centre of the cavity. These two effects will diminish the Lorentz field $\frac{4}{3}\pi I$ to some smaller value.

The fundamental difficulty is that the problem is treated as static by both Lorentz and Onsager. Lorentz assumes (implicitly) that the central dipole changes its position so rapidly that it never has time to affect the polarization of the remainder of the medium, while Onsager makes the assumption that the central dipole remains in the same direction long enough for the remainder of the medium to settle down, and for the 'back-action' of the dipole on the remainder of the medium to be treated statically. In actual fact, the two time intervals involved are probably of the same order of magnitude, so that the truth lies somewhere between the two extreme cases, and it is difficult to say which treatment should be nearer the truth. Debye's (1938) treatment using the Lorentz field seems to account for the facts satisfactorily, and it seems premature to interfere with it, though the whole question of the "local field" is very obscure, and any results obtained using this conception must be received with great caution.

We should like to thank Professor R. H. Fowler and Mr A. H. Wilson for their interest in this work. One of us (H.N V.T.) would also like to thank the Provost and Fellows of King's College, Cambridge, for the award of a studentship during the tenure of which this work was carried out.

REFERENCES

Ashmead, J. Unpublished.
Debye, P. 1938 *Ann Phys., Lpz.*, **32**, 85
Kurti, N., Laine, P. and Simon F. 1937 *C.R. Acad. Sci., Paris*, **204**, 754.
Sauer, J. 1940 *Phys. Rev.* **57**, 142.
Shire, E. S. and Barkla, H. M. 1939 *Proc. Camb. Phil. Soc.* **35**, 327.
Temperley, H. N V. 1940 *Proc. Camb. Phil. Soc.* **36**, 79.
Van Vleck, J. H 1937 *J. Chem. Phys.* **5**, 320.

Operator calculus in the electron theory of metals

By K. Fuchs

Carnegie Research Fellow, University of Edinburgh

(*Communicated by M. Born, F.R S.—Received* 28 *March* 1940)

An operator calculus is developed applicable to problems in the electron theory of metals. It differs from the common operator calculus of the quantum theory in the fact that the wave function is defined in a finite space (the atomic polyhedron) bounded by a finite surface. This leads to the introduction of surface operators.

The position operator $\mathbf{x}$ cannot be developed with respect to the proper functions of the Hamiltonian. Instead an operator $\boldsymbol{\xi}$ is introduced, which is essentially the Fourier development of $\mathbf{x}$

Thus there are three fundamental types of operators. the differential operator $\mathbf{p}$, the multiplication operator $\boldsymbol{\xi}$ and the surface operators. It is shown that with the help of these a consistent calculus can be developed

1. Introduction

The problem of metallic constitution consists principally in the solution of the wave equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = W\psi, \tag{1·1}$$

where the potential V and the wave function ψ extend throughout the whole lattice. This problem has in principle been solved by Bloch (1928), who showed that the wave functions are modulated plane waves

$$\psi_k(\mathbf{r}) = e^{i(\mathbf{k}.\mathbf{r})}u_k(\mathbf{r}), \tag{1·2}$$

where $u_k(\mathbf{r})$ has the periodicity of the lattice, and by Wigner and Seitz (1933), who noticed that Bloch's theorem reduced the problem of finding a wave function extending throughout the whole lattice, to that of finding a wave function in one atomic polyhedron, satisfying certain boundary conditions on the surface of the polyhedron They gave the first actual numerical calculation of metallic wave functions.

These investigations have revealed the character of the metallic binding. But it is rather surprising that they have not led as yet to the development of a general crystal theory of metals, comparable to the crystal theory of central forces developed by Born and others.

One reason for this—it seems to me—is to be found in the fact that the solution of the wave equation (1·1) has mostly been tackled by numerical methods. Numerical calculations are indeed to some extent indispensable, since the potential in which the electron moves is mostly determined by the nearest ion; and the potential of an ion is available only in numerical form.

However, if the solutions of the wave equation are known for a certain state of the crystal—say the undeformed crystal at the absolute zero of temperature—it should be possible to calculate the physical quantities relating to other states of the crystal (elastic constants, thermal frequencies, etc.) by means of analytical methods. An attempt in this direction has been made by Fröhlich (1937, 1938). But Fröhlich's 'perturbation of the boundary condition' can only be applied to problems for which the boundary conditions are explicitly known; this is the case only in exceptional conditions (uniform pressure, certain modes of vibration).

Furthermore, any change in the physical conditions is in general not expressible in terms of a change of boundary conditions only, but is accompanied at the same time by a change in the potential. A method including both is therefore desirable.

The ordinary perturbation theory of quantum mechanics has been applied by Wigner and Seitz (1934) for the calculation of the 'Fermi energy', due to the higher electronic states In applying this method in general certain modifications are necessary. These are due to the fact that we are dealing with a wave function defined in a finite space, the atomic polyhedron, bounded by a finite surface. This paper is devoted to a systematic investigation of the modifications necessary in order to adapt the perturbation theory of quantum mechanics to this problem. It will be seen that a consistent perturbation theory can be developed embodying the advantages of Fröhlich's method without the limitations imposed by the concentration on the boundary conditions

The essential point is the following. Let us consider the ensemble of admissible wave functions—i e. the ensemble of functions for which the Hamiltonian operator (1·1) is Hermitian. Then all operators may be admitted which transform an admissible function again into an admissible function. Now, of the two fundamental operators of quantum mechanics, the multiplication operator $\mathbf{x}$ and the differential operator $\mathbf{p} = \hbar/i\, \partial/\partial\mathbf{x}$, only the latter is admissible. The former changes an admissible wave function into another function for which the Hamiltonian is not Hermitian.

This has the consequence that if $f(\mathbf{x})$ is a function which can be developed with respect to the proper functions of the Hamiltonian operator, then $\mathbf{x}f(\mathbf{x})$ cannot be developed in this way.

I shall introduce instead of the multiplication operator another admissible operator, which coincides with **x** throughout the atomic polyhedron, except at the surface, where it is an improper operator (similar to Dirac's δ-function). Indeed, this operator is nothing but the Fourier development of **x**. The result of this replacement is that in a number of quantum mechanical operations (for example commutation with **p**) additional terms appear, which can be expressed as integrals over the surface of the atomic polyhedron. These additional terms can be regarded as the elements of an independent operator.

The main part of this paper contains the development of the general theory. In the last sections the problem of the Fermi energy is treated as an example. The reader familiar with the proof given by Wigner and Seitz will recognize it in its new form, which shows the real significance of that calculation, and holds under more general conditions.

The application to the problem of the elastic constants will be given in a following paper.

2. THE WAVE EQUATION IN THE ATOMIC POLYHEDRON

If the wave function (1·2) is inserted into the wave equation (1·1), one finds

$$H_k u_k = E_k u_k, \tag{2·1}$$

where†

$$E_k = W - \frac{\hbar^2}{2m} k^2, \tag{2·2}$$

$$H_k = \frac{p^2}{2m} + \frac{\hbar}{m}(\mathbf{k}.\mathbf{p}) + V, \quad \mathbf{p} = \frac{\hbar}{i}\nabla. \tag{2·3}$$

The function u_k has the periodicity of the lattice, and it is therefore sufficient to solve the equation (2·1) in the atomic polyhedron, subject to the boundary condition that the wave function is continuous, if periodically continued. Such functions will be called 'continuous' for short.

The 'atomic polyhedron' can of course be defined in various ways. For the present it will not be specified any further beyond the statement that the arrangement of ions throughout the whole crystal is obtained by a periodic continuation of the atomic polyhedron. The case of composite lattices, containing more than one ion per atomic polyhedron, is therefore included in the general formalism. The only property of the atomic polyhedron needed for the present is that to each point at any part of the surface there corresponds an 'opposite' point on a parallel part of the surface, the radius vector between the two points being equal to a lattice vector.

---

† For simplicity E_k will be referred to as the energy, though the 'true' energy is W.

Obviously a continuous function has the same value at two opposite points.

It can now easily be shown that the operator $\mathbf{p}$ is Hermitian for all continuous functions with continuous derivatives.

$$(u\,.\,\mathbf{p}v) - (\mathbf{p}u\,.\,v) = 0. \tag{2·4}$$

Here the scalar product $(u\,.\,v)$ is defined as the integral over an atomic polyhedron

$$(u\,.\,v) = \int u^* v\, d\tau, \tag{2·5}$$

so that (2·4) reduces to

$$\frac{\hbar}{i}\int \{u^*(\nabla v) + (\nabla u^*)\, v\}\, d\tau = \frac{\hbar}{i}\int \nabla(u^* v)\, d\tau.$$

This expression is identical with the surface integral

$$\frac{\hbar}{i}\int_S \boldsymbol{\nu} u^* v\, d\sigma,$$

where $\boldsymbol{\nu}$ is a unit vector in the direction of the outward normal. It has opposite signs at two opposite points and therefore the integral vanishes. This proves (2·4).

Now the Hamiltonian (2·3) contains $\mathbf{p}$ and p^2, apart from the potential V, which is of course Hermitian Therefore H is Hermitian for all continuous functions with continuous first and second derivatives.

In the following the usual assumption is made that the derivatives of the wave functions exist as far as they are needed and that they are continuous if the wave function is continuous. This assumption shall be implicit in the term 'continuous' function.

The solutions $u_{\underline{k}}$ of the wave equation must of course be continuous, and any operator which does not destroy the continuity is admissible. These will be called 'continuous operators' They include for example the momentum operator $\mathbf{p}$ and the potential $V(\mathbf{x})$, since the latter is itself a continuous function. (For this reason the continuity of the wave function u_k implies also the continuity of the derivatives of the wave function) However, I shall further admit operators, which destroy the continuity of the wave function, as long as the result of the operation is a function, for which the Hamiltonian is Hermitian. They can be defined as the limit of a series of continuous operators Only if the operator destroys the Hermitian character of the Hamiltonian, it shall not be admitted.

With these precautions the general operator calculus of the quantum theory can be applied. The proper solutions u_k^α of the wave equation (2·1) form a complete system of orthogonal wave functions for any given value of k, they are assumed to be normalized in the atomic polyhedron. In general it will be most convenient to select from this infinite number of systems, the system for $k=0$, which will be denoted by u_0^α. The general considerations of the following sections apply to any system and for simplicity the index k will be omitted.

By admitting in the operator calculus operators which are not continuous at the boundary, the question arises whether the perturbation theory based thereon will lead to continuous wave functions. This is indeed the case. For, by restricting the operators as outlined above, we ensure that the wave equation is satisfied not only inside the atomic polyhedron but also at the surface. Since the solutions of the wave equation are per force differentiable, they must therefore also be continuous at the boundary.

3 THE 'POSITION' OPERATOR

We have seen in the preceding section that no difficulties arise in connexion with the momentum operator **p**. This is different if we turn to the position operator **x**. For, it changes a continuous function into a discontinuous function, and in general destroys the Hermitian character of the Hamiltonian, as will be shown immediately.

Let x_i be the co-ordinates of a point inside the atomic polyhedron counted from an arbitrary origin. Let u, v be two continuous functions and operate with x_i on v. The result is a function, for which **p** is not Hermitian

$$(u\,.\,\mathbf{p}x_i v) - (\mathbf{p}u\,.\,x_i v) \neq 0.$$

For this expression is equal to

$$\frac{\hbar}{i}\int \nabla(u^* x_i v)\,d\tau = \frac{\hbar}{i}\int_s \boldsymbol{\nu}(u^* x_i v)\,d\sigma.$$

Here the right-hand side is a surface integral, which in general is different from zero. It follows therefore that x_i is not an admissible operator.

We introduce now the operator ξ_i defined by its matrix elements

$$\xi_i^{\alpha\beta} = (u^\alpha . x_i u^\beta). \qquad (3·1)$$

If x_i were an admissible operator, ξ_i would of course be identical with x_i. But since x_i is not admissible, ξ_i coincides with x_i only inside the atomic polyhedron, but not at the boundary. This is immediately evident, if we represent ξ_i as the limit of a series of operators, using the successive approximations by means of the Fourier development of x_i. Each term in this series is a continuous function of $\mathbf{x}$ and therefore an admissible operator, coinciding with the operator defined by its matrix elements. In the limit it coincides with ξ_i. Thus we have

$$(u\,.\,\mathbf{p}\boldsymbol{\xi}v)=(\mathbf{p}u\,.\,\boldsymbol{\xi}v),\quad (u\ \ H\boldsymbol{\xi}v)=(Hu\,.\,\boldsymbol{\xi}v). \tag{3·2}$$

It is obvious that $\boldsymbol{\xi}$ is Hermitian

$$(u\,.\,\boldsymbol{\xi}v)=(\boldsymbol{\xi}u\,.\,v). \tag{3·3}$$

Since $\boldsymbol{\xi}$ differs from $\mathbf{x}$ only at the boundary, all matrix elements containing $\boldsymbol{\xi}$ but no derivatives of $\boldsymbol{\xi}$ can be calculated by replacing $\boldsymbol{\xi}$ by $\mathbf{x}$ Thus

$$(\boldsymbol{\xi}\mathbf{p})^{\alpha\beta}=(u^{\alpha}\,.\,\mathbf{x}\mathbf{p}u^{\beta}), \tag{3·4}$$

but in general $\qquad (\mathbf{p}\boldsymbol{\xi})^{\alpha\beta}\neq(u^{\alpha}\,.\,\mathbf{p}\mathbf{x}u^{\beta}).$

This matrix element must be reduced to the element (3·4)

$$(\mathbf{p}\boldsymbol{\xi})^{\alpha\beta}=(\boldsymbol{\xi}\mathbf{p})^{\beta\alpha*}=(u^{\beta}\,.\,\mathbf{x}\mathbf{p}u^{\alpha})^{*}=(\mathbf{x}\mathbf{p}u^{\alpha}\ \ u^{\beta}), \tag{3·5}$$

since $\mathbf{p}$ and $\boldsymbol{\xi}$ are both Hermitian.

The components of $\boldsymbol{\xi}$ can be represented as functions of $\mathbf{x}$, therefore they commute with each other.

$$[\xi_i,\xi_j]=0. \tag{3·6}$$

4 COMMUTATION RULES AND SURFACE OPERATORS

Consider now the commutation of $\mathbf{p}$ with a function of $\boldsymbol{\xi}$. Corresponding to (3·4) and (3·5) the matrix elements are

$$\begin{aligned}[p_i,f(\boldsymbol{\xi})]^{\alpha\beta}&=-\frac{\hbar}{i}\int\left\{\frac{\partial u^{\alpha*}}{\partial x_i}f(\mathbf{x})\,u^{\beta}+u^{\alpha*}f(\mathbf{x})\frac{\partial u^{\beta}}{\partial x_i}\right\}d\tau\\&=-\frac{\hbar}{i}\int\frac{\partial}{\partial x_i}(u^{\alpha*}f(\mathbf{x})\,u^{\beta})\,d\tau+\frac{\hbar}{i}\int u^{\alpha*}\frac{\partial f(\mathbf{x})}{\partial x_i}u^{\beta}d\tau.\end{aligned}$$

The first integral can be changed into a surface integral, the second gives the matrix elements of $\partial f(\boldsymbol{\xi})/\partial\xi_i$; thus

$$[p_i,f(\boldsymbol{\xi})]^{\alpha\beta}=\frac{\hbar}{i}\left\{\frac{\partial f(\boldsymbol{\xi})}{\partial\xi_i}\right\}^{\alpha\beta}-\frac{\hbar}{i}\int_s u^{\alpha*}\nu_i f(\mathbf{x})\,u^{\beta}d\sigma. \tag{4·1}$$

The appearance of such surface integrals is characteristic for the operator calculus developed here. They can be considered as the elements of an independent operator, since all other elements in (4·1) are the elements of admissible operators. I introduce therefore the 'surface operators'

$$\nu_i(A, f(\mathbf{x}), B).$$

They are defined by means of their matrix elements

$$\nu_i(A, f(\mathbf{x}), B)^{\alpha\beta} = \int_s (A^\dagger u^\alpha)^* \nu_i f(\mathbf{x})\, B u^\beta d\sigma. \tag{4·2}$$

Here A and B are supposed to be continuous operators, so that they have no divergencies at the surface. With this definition (4·1) may be written as an operator equation

$$[p_i, f(\boldsymbol{\xi})] = \frac{\hbar}{i}\left\{\frac{\partial f(\boldsymbol{\xi})}{\partial \xi_i} - \nu_i(1, f(\mathbf{x}), 1)\right\}. \tag{4·3}$$

The surface operators can also be represented in x-space where they are improper functions. This will be seen, if the matrix elements are formed according to the usual rule and compared with the definition (4·2),

$$\nu_i(A, f(\mathbf{x}), B)^{\alpha\beta} = \int u^{\alpha *} \nu_i(A, f(\mathbf{x}), B)\, u^\beta d\tau = \int_s (A^\dagger u^\alpha)^* \nu_i f(\mathbf{x})\, B u^\beta d\sigma.$$

This equation shows that the operators $\nu_i(A, f(\mathbf{x}), B)$ are a kind of δ-function at the boundary and vanish inside the atomic polyhedron.

5. SOME PROPERTIES OF THE SURFACE OPERATORS

The surface operators have some simple properties, which will now be briefly considered.

(i) *The adjoint of a surface operator* From the definition (4·2) follows immediately that

$$\nu_i(A, f(\mathbf{x}), B)^\dagger = \nu_i(B^\dagger, f^*(\mathbf{x}), A^\dagger). \tag{5·1}$$

(ii) *Continuous functions* $f(\mathbf{x})$ If $f(\mathbf{x})$ is a continuous function of $\mathbf{x}$, it is obvious that the surface operator $\nu_i(A, f(\mathbf{x}), B)$ vanishes identically. This means that in our calculus we need not distinguish between the functions $f(\mathbf{x})$ and $f(\boldsymbol{\xi})$, if f is continuous.

(iii) *Product with continuous operators.* The most important property of the surface operators is, that a product with any continuous operator is again a surface operator In order to prove this, let us multiply the equation (4·3) from both sides with a continuous operator

$$A[p_i, f(\boldsymbol{\xi})]\, B = \frac{\hbar}{i} A \frac{\partial f(\boldsymbol{\xi})}{\partial \xi_i} B - \frac{\hbar}{i} A \nu_i(1, f(\mathbf{x}), 1)\, B.$$

The matrix elements of this equation are

$$(A\nu_i(1,f(\mathbf{x}),1)\,B)^{\alpha\beta} = \left(u^\alpha . A\frac{\partial f(\boldsymbol{\xi})}{\partial \xi_i}Bu^\beta\right) - \frac{i}{\hbar}(u^\alpha . A\{p_i f(\boldsymbol{\xi}) - f(\boldsymbol{\xi})\,p_i\}\,Bu^\beta)$$

$$= \left(A^\dagger u^\alpha . \frac{\partial f(\mathbf{x})}{\partial x_i}Bu^\beta\right) - \frac{i}{\hbar}(p_i A^\dagger u^\alpha . f(\mathbf{x})\,Bu^\beta) + \frac{i}{\hbar}(A^\dagger u^\alpha . f(\mathbf{x})\,p_i Bu^\beta).$$

If these matrix elements are written as integrals, the integrand easily reduces to a complete differential, so that the integral reduces to a surface integral, the latter is identical with the surface integral (4·2) Therefore the following relation holds

$$A\nu_i(1,f(\mathbf{x}),1)\,B = \nu_i(A,f(\mathbf{x}),B). \tag{5·2}$$

Observing that the product of two continuous operators is again a continuous operator, it follows from (5·2) that

$$C\nu_i(A,f(\mathbf{x}),B)\,D = \nu_i(CA,f(\mathbf{x}),BD), \tag{5·3}$$

where C and D are also continuous operators.

It may be remarked that if A or B in (5·2) are continuous functions of $\mathbf{x}$, they may be shifted into the middle position as follows

$$A(\mathbf{x})\,\nu_i(1,f(\mathbf{x}),1)\,B(\mathbf{x}) = \nu_i(1,A^*(\mathbf{x})f(\mathbf{x})\,B(\mathbf{x}),1). \tag{5·4}$$

This follows from the definition (4·2)

6. COMMUTATION WITH THE HAMILTONIAN AND PERTURBATION CALCULUS

Let us consider now the commutation of the Hamiltonian with a function of the ξ's. In the following the summation convention is used for the lower indices i, j, etc., denoting components of vectors. With the Hamiltonian (2·3) one finds

$$[H,f(\boldsymbol{\xi})] = \frac{p_l}{2m}[p_l,f(\boldsymbol{\xi})] + [p_l,f(\boldsymbol{\xi})]\frac{p_l}{2m} + \frac{\hbar}{m}k_l[p_l,f(\boldsymbol{\xi})] + [V,f(\boldsymbol{\xi})].$$

The last term vanishes since V and $\boldsymbol{\xi}$ are functions of $\mathbf{x}$ only. With (4·3) and (5·2) the remainder reduces to

$$[H,f(\boldsymbol{\xi})] = \frac{\hbar}{2mi}\left\{p_l\frac{\partial f(\boldsymbol{\xi})}{\partial \xi_l} + \frac{\partial f(\boldsymbol{\xi})}{\partial \xi_l}p_l + 2hk_l\frac{\partial f(\boldsymbol{\xi})}{\partial \xi_l}\right.$$
$$\left. - \nu_l(p_l,f(\mathbf{x}),1) - \nu_l(1,f(\mathbf{x}),p_l) - 2\hbar k_l\nu_l(1,f(\mathbf{x}),1)\right\}, \tag{6·1}$$

or on commuting p_l and $\partial f(\boldsymbol{\xi})/\partial \xi_l$

$$\begin{aligned}[H, f(\boldsymbol{\xi})] = \frac{\hbar}{2mi}\Big\{&2p_l\frac{\partial f(\boldsymbol{\xi})}{\partial \xi_l} + 2\hbar k_l\frac{\partial f(\boldsymbol{\xi})}{\partial \xi_l} - \nu_l(p_l, f(\mathbf{x}), 1) \\ &- \nu_l(1, f(\mathbf{x}), p_l) - 2\hbar k_l \nu_l(1, f(\mathbf{x}), 1)\Big\} \\ &+ \frac{\hbar^2}{2m}\left\{\frac{\partial^2 f(\boldsymbol{\xi})}{\partial \xi_l \partial \xi_l} - \nu_l\left(1, \frac{\partial f(\mathbf{x})}{\partial x_l}, 1\right)\right\},\end{aligned} \tag{6·2}$$

$$\begin{aligned}[H, f(\boldsymbol{\xi})] = \frac{\hbar}{2mi}\Big\{&2\frac{\partial f(\boldsymbol{\xi})}{\partial \xi_l}p_l + 2\hbar k_l\frac{\partial f(\boldsymbol{\xi})}{\partial \xi_l} - \nu_l(p_l, f(\mathbf{x}), 1) \\ &- \nu_l(1, f(\mathbf{x}), p_l) - 2\hbar k_l \nu_l(1, f(\mathbf{x}), 1)\Big\} \\ &- \frac{\hbar^2}{2m}\left\{\frac{\partial^2 f(\boldsymbol{\xi})}{\partial \xi_l \partial \xi_l} - \nu_l\left(1, \frac{\partial f(\mathbf{x})}{\partial x_l}, 1\right)\right\}.\end{aligned} \tag{6·3}$$

The commutation equations (6·1)–(6·3) are the essential equations for any perturbation problem. Their use depends largely on the special problem and I shall therefore only indicate it shortly and then treat the problem of the Fermi energy as an example.

Any change of the boundary can of course be compensated by a change of co-ordinates. It is therefore sufficient to consider a perturbation given by a perturbation Hamiltonian

$$H = \overset{0}{H} + \overset{1}{H} + \dots. \tag{6·4}$$

It will be most convenient to use the abstract perturbation calculus of Born and Jordan (1930). Our problem is then to find a unitary transformation U which brings the Hamiltonian H into the diagonal form E

$$U^{\dagger}HU = E, \tag{6·5}$$

$$U^{\dagger}U = UU^{\dagger} = 1. \tag{6·6}$$

On expansion we find the first order equation

$$[\overset{0}{H}, \overset{1}{U}] = \overset{1}{E} - \overset{1}{H}. \tag{6·7}$$

In applying this equation to problems in the theory of metals, it is often possible to find a strict solution of this equation, if the atoms are at infinite distances from each other. Let this solution be denoted by an index ∞; then one has

$$[\overset{0}{H}, \overset{1}{U}_{\infty}] = \overset{1}{E}_{\infty} - \overset{1}{H}_{\infty}.$$

Now, for infinite lattice distance the surface operators vanish, since then the wave functions vanish at the surface. For a finite lattice distance the left-hand side of this equation then differs exactly by the surface operators appearing in (6·1)–(6·3). Let these be denoted by $\overset{1}{N}$, then we have

$$[\overset{0}{H}, \overset{1}{U}_\infty] = \overset{1}{E}_\infty - \overset{1}{H}_\infty - \overset{1}{N}, \tag{6·8}$$

where $\overset{1}{U}_\infty$ and therefore $\overset{1}{N}$ are known operators. Writing

$$\overset{1}{U} = \overset{1}{U}_\infty + \overset{1}{T}, \tag{6·9}$$

and subtracting (6·8) from (6·7), we find

$$[\overset{0}{H}, \overset{1}{T}] = \overset{1}{E} - \overset{1}{E}_\infty - \overset{1}{H} + \overset{1}{H}_\infty + \overset{1}{N}. \tag{6·10}$$

The diagonal elements of the left-hand side vanish, hence

$$\overset{1}{E}{}^{\alpha} = \overset{1}{E}{}^{\alpha}_{\infty} + (\overset{1}{H} - \overset{1}{H}_\infty)^{\alpha\alpha} - \overset{1}{N}{}^{\alpha\alpha}, \tag{6·11}$$

where $\overset{1}{E}{}^{\alpha}$ is the change of the energy of the state represented by the wave function u^α. The difference $\overset{1}{H} - \overset{1}{H}_\infty$ arises only from the potential due to the ions and electrons outside the atomic polyhedron and the diagonal element can easily be approximated by classical methods. The remainder of the energy change is given by the surface integral $\overset{1}{N}{}^{\alpha\alpha}$.

The second order equation is more complicated, but in principle the same.

The equation (6·8) is of some interest. If it is remembered that $\overset{1}{N}$ is a surface operator which vanishes everywhere inside the atomic polyhedron, and disregarding the term $\overset{1}{H} - \overset{1}{H}_\infty$, it will be seen that $\overset{1}{U}_\infty$ is a solution of the equation (6·7) except at the surface Hence the corresponding wave functions satisfy the wave equation everywhere inside the polyhedron, but not at the boundary. This is due to the fact that the operator $\overset{1}{U}_\infty$ is not a continuous operator, and as pointed out in § 2, for this reason the corresponding wave function is automatically rejected

7 THE WAVE FUNCTIONS AND ENERGIES OF THE HIGHER ELECTRONIC STATES

In this section let us consider a cubic lattice and assume that the wave functions and energies for $k = 0$ are known. These form a complete orthogonal system of wave functions and the solution of the wave equation for the higher states ($k \neq 0$) can be obtained by the perturbation theory.

The results for the energy have already been given by Wigner and Seitz (1934). But they have not paid any explicit attention to the question whether the operations performed are admissible or not.

It will be most convenient to assume the atomic polyhedron in such a way that it symmetrically surrounds the lattice points. Then the Hamiltonian $\overset{0}{H}$ for $k = 0$ has the symmetry of the lattice and the proper functions $\overset{0}{u}$ are symmetric or antisymmetric with the lattice symmetry. For a cubic lattice it follows that the diagonal elements of the operators $\mathbf{p}$ and $\boldsymbol{\xi}$ vanish

$$(\mathbf{p})^{\alpha\alpha} = (\boldsymbol{\xi})^{\alpha\alpha} = 0. \tag{7·1}$$

Expanding with regard to k the Hamiltonian is

$$H = \overset{0}{H} + H_k, \tag{7·2}$$

where from (2·3)

$$\overset{0}{H} = \frac{p^2}{2m} + V, \tag{7·3}$$

$$H_k = \frac{\hbar}{m}(\mathbf{k}.\mathbf{p}). \tag{7·4}$$

Let the energy be

$$E = \overset{0}{E} + E_k. \tag{7·5}$$

Then (6·5) gives on multiplication with U (observe that $\overset{0}{E} \equiv \overset{0}{H}$)

$$[\overset{0}{H}, U] + H_k U - U E_k = 0. \tag{7·6}$$

Let us consider the solution of this equation for infinite lattice distance. In this case the wave functions ψ_k inside a given zone are all identical and equal to $\psi_0 = u_0$. Therefore it follows from (1·2) that $u_k = \exp(-i(\mathbf{k}.\mathbf{x}))\, u_0$ Hence the unitary transformation U for infinite lattice distance is

$$U_\infty = e^{-i(\mathbf{k}\,\boldsymbol{\xi})}.$$

This operator is uniquely defined since the ξ_l commute with each other. We are thus led to the substitution

$$U = e^{-i(\mathbf{k}\,\boldsymbol{\xi})} + T. \tag{7·7}$$

Now, from (6·2) it follows that

$$[\overset{0}{H}, e^{-i(\mathbf{k}\,\boldsymbol{\xi})}] + \frac{\hbar}{m}(\mathbf{k}.\mathbf{p})\, e^{-i(\mathbf{k}.\boldsymbol{\xi})} + \frac{\hbar^2}{2m} k^2 e^{-i(\mathbf{k}.\boldsymbol{\xi})} - N = 0, \tag{7·8}$$

where N is the surface operator

$$N = \frac{i\hbar}{2m}\{\nu_l(p_l, e^{-i(\mathbf{k}\,\mathbf{x})}, 1) + \nu_l(1, e^{-i(\mathbf{k}\,\mathbf{x})}, p_l) + \hbar k_l \nu_l(1, e^{-i(\mathbf{k}\,\mathbf{x})}, 1)\}. \tag{7·9}$$

Subtracting (7·8) from (7·5), one finds with (7·4)

$$[\overset{0}{H}, T] + H_k T - T E_k - e^{-i(\mathbf{k}\,\boldsymbol{\xi})} W_k + N = 0. \tag{7·10}$$

Here

$$W_k = E_k + \frac{\hbar^2}{2m} k^2 \tag{7 11}$$

is the true energy (cf (2·2))

In order to solve this equation, we expand it in powers of k Let be

$$\left.\begin{aligned} H_k &= \overset{1}{H} = \frac{\hbar}{m}(\mathbf{k}.\mathbf{p}), \quad & T &= \overset{1}{T} + \overset{2}{T} + . \quad , \\ W_k &= \overset{1}{W} + \overset{2}{W} + \ldots, \quad & N &= \overset{1}{N} + \overset{2}{N} + \quad , \end{aligned}\right\} \tag{7·12}$$

where

$$\overset{1}{N} = \frac{\hbar}{2m}\{\nu_l(p_l, (\mathbf{k}\ \mathbf{x}), 1) + \nu_l(1, (\mathbf{k}\ \mathbf{x}), p_l)\}, \tag{7·13}$$

$$\overset{2}{N} = -\frac{i\hbar}{4m}\{\nu_l(p_l, (\mathbf{k}\ \mathbf{x})^2, 1) + \nu_l(1, (\mathbf{k}\ \mathbf{x})^2, p_l) + 2i\hbar k_l \nu_l(1, (\mathbf{k}.\mathbf{x}), 1)\} \tag{7·14}$$

Substitution of these expansions in (7·10) gives the first and second order equations

$$[\overset{0}{H}, \overset{1}{T}] - \overset{1}{W} + \overset{1}{N} = 0, \tag{7·15}$$

$$[\overset{0}{H}, \overset{2}{T}] + \overset{1}{H}\overset{1}{T} - \overset{1}{T}\overset{1}{W} - \overset{2}{W} + i(\mathbf{k}.\boldsymbol{\xi})\,\overset{1}{W} + \overset{2}{N} = 0. \tag{7·16}$$

Further, from the condition (6·6) follows

$$\overset{1}{T} + \overset{1}{T}{}^{\dagger} = 0, \tag{7·17}$$

$$\overset{2}{T} + \overset{2}{T}{}^{\dagger} + i(\mathbf{k}.\boldsymbol{\xi})\,\overset{1}{T} - \overset{1}{T}{}^{\dagger}\,i(\mathbf{k}\ \boldsymbol{\xi}) + \overset{1}{T}\overset{1}{T}{}^{\dagger} = 0. \tag{7·18}$$

From the definition (4·2) of the elements of a surface operator follows that both terms (7·13) in the diagonal elements of $\overset{1}{N}$ vanish, owing to the cubic symmetry. Therefore the diagonal elements of (7·15) give

$$\overset{1}{W}{}^{\alpha} = \overset{1}{N}{}^{\alpha\alpha} = 0. \tag{7·19}$$

The non-diagonal elements of (7·15) give

$$\overset{1}{T}{}^{\alpha\beta} = \frac{\overset{1}{N}{}^{\alpha\beta}}{\overset{0}{E}{}^{\beta} - \overset{0}{E}{}^{\alpha}}, \quad \overset{1}{T}{}^{\alpha\alpha} = 0, \tag{7·20}$$

whereas the diagonal elements of $\overset{1}{T}$ vanish in view of (7·17), since they may be taken to be real without loss of generality.

The second order equation (7·16) now reduces to

$$[\overset{0}{H}, \overset{2}{T}] + \overset{1}{H}\overset{1}{T} - \overset{2}{W} + \overset{2}{N} = 0, \tag{7·21}$$

with the solution

$$\overset{2}{W^{\alpha}} = \overset{2}{N^{\alpha\alpha}} + (\overset{1}{H}\overset{1}{T})^{\alpha\alpha}, \tag{7·22}$$

$$\overset{2}{T^{\alpha\beta}} = \frac{(\overset{1}{H}\overset{1}{T})^{\alpha\beta} + \overset{2}{N^{\alpha\beta}}}{\overset{0}{E^{\beta}} - \overset{0}{E^{\alpha}}}. \tag{7·23}$$

The diagonal element of $\overset{1}{H}\overset{1}{T}$ is with (7·20)

$$(\overset{1}{H}\overset{1}{T})^{\alpha\alpha} = \sum_{\beta} \overset{1}{H^{\alpha\beta}} \frac{\overset{1}{N^{\beta\alpha}}}{\overset{0}{E^{\alpha}} - \overset{0}{E^{\beta}}}. \tag{7·24}$$

Following Wigner and Seitz (1934) this sum can be evaluated approximately for the lowest zone of the valence electrons. For in this case all states of lower energy give a negligible contribution, since their wave functions are very small at the boundary so that the corresponding element of the surface operator $\overset{1}{N}$ is also small. We may therefore replace $\overset{0}{E^{\beta}}$ in the numerator of (7·24) by a suitably chosen mean value, which is of the order of the energy of the first state giving an appreciable contribution. Substituting for $\overset{1}{H}$ the value (7·12), the energy (7·22) is in this way

$$\overset{2}{W^{\alpha}} = \overset{2}{N^{\alpha\alpha}} + \frac{\hbar}{m} \frac{k_l(p_l\overset{1}{N})^{\alpha\alpha}}{\overset{0}{E^{\alpha}} - \overset{0}{\bar{E}}}, \tag{7·25}$$

or with (7·13), (7·14) and (5·2)

$$\begin{aligned}\overset{2}{W^{\alpha}} = -\frac{i\hbar}{4m}\{\nu_l(p_l, (\mathbf{k}.\mathbf{x})^2, 1) + \nu_l(1, (\mathbf{k}.\mathbf{x})^2, p_l) + 2i\hbar k_l \nu_l(1, (\mathbf{k}.\mathbf{x}), 1)\}^{\alpha\alpha} \\ + \frac{h^2}{2m^2} \frac{k_l}{\overset{0}{E^{\alpha}} - \overset{0}{\bar{E}}} \{\nu_i(p_l\, p_i, (\mathbf{k}.\mathbf{x}), 1) + \nu_i(p_l, (\mathbf{k}.\mathbf{x}), p_i\}^{\alpha\alpha}.\end{aligned} \tag{7·26}$$

Thus the energy is given in terms of surface integrals containing only the wave function u_0^{α} of the lowest valence state. If the latter is an *s*-state in the undeformed cubic lattice, most terms vanish in view of the cubic symmetry or the boundary condition

$$\frac{\partial u_0^{\alpha}}{\partial \nu} = 0.$$

It then follows that

$$\overset{2}{W}{}^{\alpha} = \frac{\hbar^2}{2m} k^2 \left\{ \nu_x(1, x, 1)^{\alpha\alpha} + \frac{1}{\overset{0}{E}{}^{\alpha} - \overset{0}{E}} \frac{\nu_l(p_l p_x, x, 1)^{\alpha\alpha}}{m} \right\}. \qquad (7\cdot27)$$

Here the index x is used so as to avoid conflict with the summation convention.

In the spherical approximation this expression gives the value found by Wigner and Seitz (1934)

$$\overset{2}{W}{}^{\alpha} = \frac{\hbar^2}{2m} k^2 f^{\alpha} \left(1 + 2 \frac{V(r_0) - \overset{0}{E}{}^{\alpha}}{\overset{0}{E} - \overset{0}{E}{}^{\alpha}} \right), \quad f^{\alpha} = \frac{4\pi}{3} r_0^3 (u^{\alpha} u^{\alpha *})_{r=r_0}. \qquad (7\cdot28)$$

Numerical calculations for sodium (Wigner and Seitz 1934) and copper (Fuchs 1935) have shown that the second term in (7·28) is small. A collection of numerical values of f^{α} is given by Mott and Jones (1936).

This calculation can readily be extended to the case of metals with other than cubic symmetry. In this case

$$\overset{1}{W}{}^{\alpha} = \overset{1}{N}{}^{\alpha\alpha} \qquad (7\cdot29)$$

does not necessarily vanish. Observing that $\overset{1}{H}$, $\overset{1}{W}$ and $\overset{2}{W}$ are Hermitian, (7·16) gives with (7·17)

$$\overset{2}{W} = \tfrac{1}{2}\{\overset{1}{H}\overset{1}{T} - \overset{1}{T}\overset{1}{H} + \overset{2}{N} + \overset{2}{N}{}^{\dagger} + [\overset{1}{W}, \overset{1}{T} - i(\mathbf{k}.\boldsymbol{\xi})] + [\overset{0}{H}, \overset{2}{T}] + [\overset{2}{T}{}^{\dagger}, \overset{0}{H}]\} \qquad (7\cdot30)$$

The commutators in this equation have vanishing diagonal terms, since $\overset{0}{H}$ and $\overset{1}{W}$ are diagonal Therefore one has

$$\overset{2}{W}{}^{\alpha} = \tfrac{1}{2}\{\overset{1}{H}\overset{1}{T} - \overset{1}{T}\overset{1}{H} + \overset{2}{N} + \overset{2}{N}{}^{\dagger}\}^{\alpha\alpha}. \qquad (7\cdot31)$$

The diagonal term of $\overset{1}{T}\overset{1}{H}$ can be evaluated in the same way as that of $\overset{1}{H}\overset{1}{T}$ with the result

$$\tfrac{1}{2}(\overset{1}{H}\overset{1}{T} - \overset{1}{T}\overset{1}{H})^{\alpha\alpha} = \frac{\hbar}{2m} \frac{k_l}{\overset{0}{E}{}^{\alpha} - \overset{0}{E}} \{p_l \overset{1}{N} + \overset{1}{N} p_l\}^{\alpha\alpha} - \frac{1}{\overset{0}{E}{}^{\alpha} - \overset{0}{E}} \frac{\hbar}{m} (\mathbf{k}.\mathbf{p})^{\alpha\alpha}.\overset{1}{N}{}^{\alpha\alpha}. \qquad (7\cdot32)$$

Here the last term arises from the fact that the term $\beta = \alpha$ in the summation (7·24) must be omitted; it is zero for the cubic lattice.

If we apply the common perturbation theory directly, we have of course the well-known equation ($\overset{1}{W}{}^{\alpha}$ is identical with $\overset{1}{E}{}^{\alpha}$)

$$\overset{1}{W}{}^{\alpha} = \overset{1}{H}{}^{\alpha\alpha} = \frac{\hbar}{m}(\mathbf{k}.\mathbf{p})^{\alpha\alpha}. \tag{7·33}$$

Therefore with (7·29)

$$\frac{\hbar}{m}(\mathbf{k}.\mathbf{p})^{\alpha\alpha} = \overset{1}{N}{}^{\alpha\alpha} = \overset{1}{W}{}^{\alpha}, \tag{7·34}$$

so that the last term in (7·32) is known.

Further, it follows from (7·14) and (5·1) that

$$\tfrac{1}{2}(\overset{2}{N} + \overset{2}{N}{}^{\dagger}) = \frac{\hbar^2}{2m} k_l \nu_l(1, (\mathbf{k}.\mathbf{x}), 1). \tag{7·35}$$

Thus we have for any lattice

$$\overset{2}{W}{}^{\alpha} = \frac{\hbar^2}{2m} k_l \nu_l(1, (\mathbf{k}.\mathbf{x}), 1)^{\alpha\alpha} + \frac{1}{\overset{0}{E}{}^{\alpha} - \overset{\bar{0}}{E}} \Big[-(\overset{1}{W}{}^{\alpha})^2 + \frac{\hbar^2}{4m^4} k_l \{\nu_i(p_l p_i, (\mathbf{k}.\mathbf{x}), 1)$$
$$+ \nu_i(p_l, (\mathbf{k}.\mathbf{x}), p_i) + \nu_i(p_i, (\mathbf{k}\ \mathbf{x}), p_l) + \nu_i(1, (\mathbf{k}\ \mathbf{x}), p_i p_l\}^{\alpha\alpha} \Big]. \tag{7·36}$$

REFERENCES

Bloch, F. 1928 *Z. Phys* **55**, 555.

Born, M. and Jordan, P. 1930 *Elementare Quanten Mechanik* (see particularly pp. 194–195) Berlin Springer.

Frohlich, H. 1937 *Proc. Roy. Soc* A, **158**, 97.

— 1938 *Phys Rev* **54**, 945

Fuchs, K 1935 *Proc Roy. Soc.* A, **151**, 585

Mott, N. F and Jones, H. 1936 *The theory of the properties of metals and alloys* (see particularly p 82). Oxford Univ. Press

Wigner, E and Seitz, F. 1933 *Phys. Rev.* **43**, 804.

— — 1934 *Phys. Rev.* **46**, 509.

The soft X-ray spectroscopy of solids

II. Emission spectra from simple chemical compounds

By H. M. O'Bryan, *Georgetown University, Washington, D.C.*
and H. W. B. Skinner, *Wills Physical Laboratory, Bristol*

(*Communicated by A. M. Tyndall, F.R.S.—Received 9 April 1940*)

K- and *L*-emission bands, emitted in the soft X-ray region of the spectrum by atoms in a number of chemical compounds, have been investigated. The substances for which such spectra are available through the present or previous work are fluorides, chlorides, bromides, iodides, oxides, sulphides; boron nitride and other boron compounds, and carbides. In the cases of SiC, BN and a number of oxides, spectra from both the component atoms forming the compound have been obtained.

The results are interpreted, as those for metals and element-insulators have been in a previous paper (Skinner 1940), to give the characteristics of the bands of levels which exist for the valence-electrons in the normal state of the substance. In the case of the halides, data on the *p*- and *s*-levels of the negative ions are given, and it is shown that, even in this case, the crystal structure leaves its mark on the form of the bands of levels. The $2p$-bands of oxygen from most oxides are more spread-out on an energy scale, thus showing that the interaction between electrons in neighbouring atoms is considerable. The spectra of the metal ions in oxides are very complex, and an attempt is made to disentangle the factors which lead to this complication. In doing so, it is hoped that we may have thrown some light on the difficult subject of the structure of semi-polar compounds in solid form.

1. Introduction

In the first paper of the series, a full account has been given of the soft X-ray bands of metals and element insulators of the first two groups of the periodic table, and of their use in the formulation of a detailed knowledge of the systems of electron levels which exist in such solids (Skinner 1940, to be referred to as 'Paper I'). In the present paper, we propose to describe the results of similar investigations on chemical compounds, chiefly halides, oxides and sulphides. Though the bands are often more complicated than for the elements, we shall attempt to interpret them along lines analogous to those used in the previous paper.

A solid chemical compound of course consists of a lattice of atoms or ions of more than one sort. We therefore have the possibility of observing spectra of two or more kinds from a given compound, for example, the *K*-spectrum of Be and the *K*-spectrum of O in BeO. Though it has not been always possible to observe all these component spectra on account of technical difficulties, they have been obtained for a number of oxides and for BN. The result is that the K-spectrum of O (or N) is quite different from the corresponding K- or L-spectrum of the metal ion, the latter being extremely complex. It is also quite different from the same spectrum from the metal itself, as was first shown by Karlsson and Siegbahn (1934) for the K-spectra of Mg and Al in MgO and Al_2O_3. The same type of difference was also shown between the K-spectra of B obtained from the solid element and from B_2O_3 by Gwinner and Kiessig (1937) and by Hautot and Serpe (1937). We have added to the detail observed in the latter case, and have also investigated a number of other cases of metal-ion spectra from oxides, etc., including some L_{23}-spectra,* which show somewhat different peculiarities †

The K-spectra of O from various oxides had, in comparison, been neglected. Work published by several authors‡ had established the existence and wave-length of the spectrum, but seems to suffer from the defect of inadequate resolution or of insufficiently clean conditions in the X-ray tube. These spectra, of which we give a number, are characteristic of the particular oxide from which they are emitted, and are among the most interesting bands in the whole field of soft X-rays

In the case of the halides, we did not succeed in obtaining metal-ion spectra. Work on the L-spectra of chlorine and the M-spectra of bromine from a number of chlorides and bromides has already been published by Siegbahn and Magnusson (1934*b*, 1935). We have added some examples, particularly for the bromides, and have generally confirmed their experimental work. Our determinations of the K-spectra of the fluorides are quite new, as is also our work on the K-spectrum of nitrogen from BN, and the L_{23}-spectra of sulphur from several sulphides

We shall only be concerned in the present paper with one example which can be classed as a genuinely covalent compound. This is SiC, for which we give the K-spectrum of C and the L_{23}-spectrum of Si. Previous work has supplied a considerable amount of data on the K-spectra of C and B obtained from various compounds of this type.§

---

* When the (L_2-L_3) energy separation is small and can be disregarded, we write for convenience L_{23} instead of L_2 and L_3.

† The L_{23}-band of Al in Al_2O_3 published by Siegbahn and Magnusson (1934*a*) is not complete

‡ Hautot (1934), Magnusson (1938). After the present paper was partly written, the work of Tyrén (1939) came to hand This gives results in agreement with ours.

§ See Siegbahn and Magnusson (1935), Broili, Glocker and Kiessig (1934), and Gwinner and Kiessig (1937). Also for work on alloys of the light metals: Yoshida (1936), Skinner and Johnston (1937*b*) and Farineau (1938)

The *K*-spectra of oxygen and fluorine lie near 23·6 and 18·4 A respectively, and are thus in a region which provides a reasonable resolution of the electronic level-system under investigation (see Paper I, § 3; also Skinner 1939). The remaining spectra of this paper lie, from this point of view, in an even more favourable region. There exist also a number of investigations of hard X-rays, in which small changes of line-form are observed in passing from element to chemical compound. Owing to poor theoretical resolving power due to the short wave-length, such data are not of much use for providing material for determining the level systems of the solids in question.

2. Experimental technique

The range of wave-lengths included in the present investigations stretches from about 17 A to about 300 A. Two vacuum spectrographs of grazing-incidence type were used, each of which had a glass grating ruled with 30,000 lines per inch by Professor R. W. Wood, both were used with the photographic plate bent along the Rowland circle. One had a radius of curvature of 7 m. and was adjusted to an angle of incidence of 1·5°. This spectrograph was employed for wave-lengths below 50 A, namely, for the *K*-radiations of F, O, N and C. The other spectrograph, with a 1 m. grating at 6°, was that employed for the experimental work of Paper I. The slit widths used were generally about 1·5/100 mm. for the first spectrograph and 2/100 mm for the second These give respectively voltage resolutions of the radiation, taking into account the slit-width of the microphotometer, of about 0·7 eV at 23 A and about 0·4 eV at 100 A, for first order spectra With the 1 m. spectrograph, high-order spectra were often employed; but, unfortunately, the second-order spectra given by the 7 m. grating proved to be very weak, and most of the work was done using first orders A few second-order spectra of oxygen *K*-radiation from CaO appeared to show that the resolution in our first-order spectra was sufficient, but, in other cases, particularly the fluorine *K*-spectra, no definite check was made. The resolution in the long wave-length spectra was certainly adequate for all their features.

The construction of the spectrographs was similar to that described by Skinner and Johnston (1937*a*), and the general technique the same as in Paper I. So neither will be described in detail here. The glass X-ray tube, exhausted by oil diffusion pumps and provided with liquid-air traps, were the same as before. The substance to be investigated was deposited on one of the four sides of the square anticathode, which could be rotated so that any of the sides could be put in position for the exposure. The X-ray filament was always run with one side of the anticathode facing it for an hour or more as a preliminary. When the pressure had fallen to a low enough value, the other sides of the anticathode whose surfaces were still fresh, were used successively for the exposures, several of which could be obtained on the same plate by means of a shutter working close to the photographic emulsion. When possible the material to be used was evaporated on to the anticathode, this method was used for the halides and for ZnS. Other substances, for example oxides, were emulsified with a pure volatile liquid such as carbon tetrachloride, and deposited in a thin uniform layer on the polished copper surface of the anticathode. Experience showed the required thickness, and, under such conditions, the deposit was often practically unaffected by the bombardment of the electron current. Except for the halides, the copper remained bright during the exposures, and the results themselves show that they are free from contamination, for instance, by Cu_2O In the case of the halides, the method of evaporation made possible the frequent renewal of the material. This is very necessary, since the halides are among the substances which decompose when bombarded by electrons. Other such compounds are carbonates, sulphates and cyanides, which become changed into carbides, sulphides and oxides, giving off oxygen or nitrogen. The results obtained from such substances are therefore unreliable.* In the case of the halides, however, the decomposition takes the form of the evolution of halogen gas, leaving the pure metal The gas is pumped off, and thus there is no contamination of the halogen spectrum characteristic of the halide, due to other compounds On the other hand, the corresponding metal-ion spectra are spoiled. This kind of decomposition appears to be due to a type of electrolysis, and happens especially quickly when the positive and negative ions are small and mobile; LiF is the worst case Most of the oxides used were quite stable and lasted almost indefinitely on the target. But Al_2O_3 was found to decompose slowly, giving Al metal,† even so, it was possible to separate out the characteristic metal-ion spectrum of the compound from the metal spectrum by a knowledge of the shape of the latter. The oxides of iron, FeO and Fe_2O_3, give the same result for the K-spectrum of oxygen, and it is probable that the latter compound decomposes with the liberation of oxygen to give FeO.

---

* Siegbahn and Magnusson's (1935) results for carbon K-radiation from some carbonates must therefore be regarded with some suspicion

† Karlsson and Siegbahn (1934) publish a K-spectrum of Al in Al_2O_3 which shows no sign of decomposition. The exposure may have been short.

The 1 m. spectrograph was used under running conditions identical with those of Paper I. Currents of 10–100 mA at 3500 V were used in the X-ray tube, but the spectra of metal ions in oxides are much weaker than those emitted by the pure metals. The lower light-gathering power of the 7 m. spectrograph made the use of a larger input in the X-ray tube desirable; currents up to 150 mA at 8000 V were actually used. The pressure conditions were therefore liable to be somewhat inferior, but there was no sign that the results were adversely affected. Exposures, which ranged from a few minutes to several hours, were taken on Ilford Q_1 plates. The photometry was not so carefully done as that described in Paper I, and plate-calibration curves were not taken in all individual cases. On the other hand, the experience of Paper I and of some plate calibrations at the shorter wave-lengths, shows that no large errors will be made in taking the photometer deflexion proportional to the intensity of the radiation causing the blackening of the emulsion, at least in the region of densities covered in most of the exposures. Though a few of the plates have regions of exceptional density, for which the corresponding intensity will be underrated, no more elaborate method of conversion to intensities was carried out.

Wave-length measurements were only rough, especially in the short wave-length region, where we had no standard available, except for the soft X-ray bands from various Fe-group metals, whose wave-lengths are given by Tyrén (1937). However, the wave-length of the peak of the oxygen K-band from MgO was determined approximately by comparison of its ninth order with the emission edge of metallic Mg. The result is $23 \cdot 61 \pm 0 \cdot 02$ A * Since we had no comparison lines on the plates, our investigation of the variation of this wave-length with the nature of the oxide was quite incomplete In some cases, however, we had the results from different oxides on adjacent strips of the photographic plate and found a small variation in the peak wave-length The peak wave-length for BeO, B_2O_3, MgO, Al_2O_3 seemed to be the same within about 1/100 A, for SiO_2, it is shorter by about 3/100 A, while for CaO and SrO it is longer by about 2/100 A. In a diagram of the results to be given in figure 7, since the shift of the peak in some cases was not determined, we have arbitrarily arranged the peaks of the curves below one another in two groups, and the wave-length scales for some of the oxides are therefore slightly displaced. In the long wave-length region of the spectrum we have either taken wave-lengths from Siegbahn and Magnusson (1934*b*,

---

* Tyrén (1939) gives the more accurate value 23·603 A for MgO. His results for the oxides of Be, B, Mg, Al, Ca, Sc, Sr, Y, Ba vary from 23 56 to 23 61 A.

1935) or have made comparisons with high-order carbon K-spectra, or with metal emission edges, but again no great precision is claimed.

3. GENERAL INTERPRETATION OF RESULTS

In Paper I, § 3, we have given the general lines of interpretation of soft X-ray emission bands. The excited state is regarded as a discrete state, differing in energy from the normal state of the material by an amount W_x per atom, or molecule. After the transition giving rise to the emission of a soft X-ray quantum of frequency ν, the final state is the normal state of the material, except that an electron is missing from the band of filled valence-electron levels. For convenience, the energy level from which the electron is missing is defined as one of energy E, where E is measured from the bottom of the band of levels. We therefore characterize the energy per molecule of the material after the transition as $W(E)$ and we have

$$h\nu(E) = W_x - W(E) \tag{1}$$

It is shown in Paper I, § 3, that if we call $I(E)$ the observed intensity in a band of radiation, then

$$\left.\begin{aligned} I(E)/\nu^3 &\propto N_p(E) && \text{for a } K\text{-spectrum,} \\ &\propto N_{s+d}(E) && \text{for an } L_2\text{- or } L_3\text{-spectrum,} \end{aligned}\right\} \tag{2}$$

where $N_p(E)$, $N_{s+d}(E)$ are components of the function $N(E)$ which represents the *density* of levels at the energy E. $N_s(E)$, $N_p(E)$, ... represent those levels which, *near to the nuclei*, give rise to wave functions similar to atomic s, p, ... wave functions and

$$N_{s+d}(E) = N_s(E) + \alpha N_d(E),$$

where α is a constant which cannot be determined experimentally. In this paper, however, we shall almost always be able to take $N_d(E)$ as zero, since the atoms with which we have to deal are mostly unlikely to have any d-electrons in their normal state.

In the examples of Paper I, the quantities $N_s(E)$, $N_p(E)$ were found to have positive values within the same range of values of E, in other words, the bands of s- and p-like levels overlap completely. But in most of the cases which concern us in the present paper, the problem is simplified in that the s- and p-levels overlap only slightly, or not at all. This is due to the fact that we shall mainly be concerned with levels similar to those of

F^- in fluorides. For these ions, the $(2s-2p)$ separation in the free state is relatively large (of the order of 10 eV) and the interaction in the lattice is not sufficiently great to amalgamate the bands which correspond to these ionic levels. It follows that in such cases, the values of $N_s(E)$ and $N_p(E)$ have more direct meanings than they do for example in a metal where the admixture of s- and p-levels is complete In fact they represent the function $N(E)$ itself Thus, for example in the case of the K-spectrum of fluorine in a fluoride, we can interpret our results directly as giving $N(E)$ for the $2p$-band of the negative ions.

In the case of crystals which are not purely ionic, we shall have to bear in mind the fact that electrons regarded as attached to negative ions have a certain probability of being found in the region of the positive ions. This introduces a complication, since obviously the space symmetry of a wave function near a certain nucleus may be different from that near a different nucleus. Thus, in the case of semi-polar compounds, our functions $N_s(E)$, $N_p(E)$ have no meaning except in relation to a given kind of atoms in the lattice. Usually, it is convenient to refer then to the negative ion, and we shall speak, for example, of the $2p(O)$ band of levels in an oxide, or where no ambiguity can arise, simply to the $2p$-band of the oxide.

In the case of chemical compounds, an important phenomenon is found which plays only a small part when we are dealing with solid elements. This is the observation of *multiple emission bands*. In some cases, the features of the $I(E)$ curve repeat themselves with more or less accuracy at different values of E, in other cases, a second and different $I(E)$ curve is present which seems to have no connexion with the main $I(E)$ curve. In the case of ordinary hard X-ray spectra, a similar effect is observed in the emission of satellite lines, while in Paper I, § 8, evidence has been given for satellite bands, similar in form to the main bands of the L_{23}-spectra of Na, Mg, Al and Si in the form of solid elements. Such satellites, however, have an intensity of only about 1 % of that of the main bands, whereas in some of the cases with which we shall be dealing the satellites may have an intensity of the same order of magnitude as that of the main bands. Their existence is certainly an intrinsic property of chemical compounds and unfortunately makes the interpretation of the soft X-ray bands much more difficult than for simple substances.

All satellite bands can be understood, according to equation (1), only by postulating that, corresponding to a given type of excitation (K, L_2, L_3, etc.), we have to take into account either multiple values of W_x, of $W(E)$ or of both. If we exclude cases in which the material of the target is so modified or decomposed by the general bombardment in the X-ray tube as to give spurious emission bands, we may distinguish between two classes of satellites:

(A) A subsidiary excited state W_x^1 is produced, as well as the main excited state W_x^0 by the ionization process due to the impinging electron in the X-ray tube; or by a secondary effect after the initial ionization process. Into this class fall the satellite lines of ordinary X-rays, produced by original double ionization of the atom core; or, after an original single ionization, by the occurrence of an Auger process (see Paper I, § 8) which ionizes it again before radiation is emitted. All satellites of this class are very weak.

(B) In the case of partly ionic compounds (and this includes nearly all chemical compounds) the number of valence electrons associated with a particular atom at any instant may not be a constant for all atoms. Under these conditions, multiple emission bands may occur.

Underlying all cases is the possibility of variation in the number of electrons in the emitting atom. We shall first calculate the values of W_x for *free atoms or ions*. It is easily seen that

$$W_x = u_x - \sum^{n} (u'_n - u_n), \tag{3}$$

where u_n is the nth successive ionization potential of the atomic system, u_x is the ionization potential of the X-ray shell when all the n exterior electrons are removed, and u'_n is the nth successive ionization potential of the atom when an electron is missing from the X-ray shell. An approximation to the values u'_n may be obtained by taking them to be equal to the successive ionization potentials of the atom next higher in the periodic table to the atom considered. As an example, we may calculate, using equations (1) and (3), the difference in energy between the X-ray lines emitted by a negative ion and a neutral atom (e g the *K*-lines of F and F) We obtain

$$W_x^0 - W_x^- = u'_1 - u_1, \tag{4}$$

$$h\nu^\circ - h\nu^- = u'_1 - u_2, \tag{5}$$

where u_1 is the first ionization potential of F^- and u_2 is the first ionization potential of F. The *K*-emission line for neutral F is displaced about 3 eV towards high energies from the line for F^-. On the other hand, the change in the values of W_x is much greater, namely, about 16 eV. The difference is due to the higher value of W (equation (1)) for F than for F^-.

When we pass to the case of the solid, we must distinguish between two types of satellites, *line satellites* and *band satellites*, both of which are found in different instances. In order to understand the possibilities, we must consider in more detail the structures of the intermediate excited state, W_x, and the final state $W(E)$. The intermediate state is formed when an electron from an X-ray shell is removed by electron impact, a very rapid process requiring a time of the order of 10^{-17} sec. At this moment, the valence electrons are subjected to an increased field of force, on account of the increased core charge. There appears to be two possibilities:

(1) The system of valence electron levels may be unaffected by the altered field of force of a single nucleus, and no new valence levels appear.

(2) The valence electrons associated with a particular atom may be 'sucked in' by the increased field of force to form sharp levels of 'atomic' character.

We may call these '*non-atomic*' and '*atomic*' *intermediate states*. The atomic type of state may be expected in the case of very polar compounds, the non-atomic type of state must certainly be assumed to occur in the case of metals. This is proved by the observed band structure of the L-satellites of the second group metals (Paper I, figure 6) These are of class (A) and correspond to the creation of *two* vacancies in the L-shell, one of which persists throughout the emission process. The observed radiation therefore corresponds to values of $W(E)$ in the case when an L-electron is missing (as well as a valence electron) after the radiation is emitted. The fact that a *band* of radiation is found proves that the valence electrons in the neighbourhood of an L-ionized atom in a metal must be considered to belong to a band of levels, the same band, in fact, as exists throughout the lattice, in other words the valence levels of the L excited atom in the metal are of non-atomic type. We may probably generalize this result to all cases of soft X-ray emission from metals. The vacancy in the X-ray shell merely causes a redistribution of the energies of electrons likely to be found in the neighbourhood of the emitting atom, electrons from lower levels of the band being heavily favoured. This redistribution is very well shown in the L satellite bands by the fact that they differ from the main bands by a considerably greater intensity of radiation of the lower energies, it is especially obvious in the case of Na. The present point of view* is the same as that put forward by Skinner and Johnston (1937*b*) to explain the results obtained on the soft X-ray bands emitted from an element present as a dilute solid solution in another metal.

---

* The opposite point of view was put forward by Skinner (1932) to explain the fact that, whereas the value of W_s for the free Li atom is 75 eV, the mean energy of the K-band emitted by Li metal is only 54 eV. The above argument, however, seems conclusive; the energy difference must be explained by the redistribution of all the electrons in the neighbourhood of the X-ray excited atom towards lower levels in the valence band.

Thus our discussion of these satellites of class (A) has led to the conclusion that X-ray excited states of non-atomic character can exist, and, in fact, do exist in the case of metals. In the case of chemical compounds, we shall see that non-atomic X-ray excited states also occur, as might be expected; but in certain cases, atomic states also occur. When they happen they lead to the emission of satellites of class (B).

In a partly ionic compound, for example, a fluoride, *the number of valence electrons associated with a particular atom in the crystal is not to be regarded as a constant*, but as varying from time to time. Usually, the available valence electrons will be found in the neighbourhood of the fluorine centres, which will therefore have the character of negative ions. Occasionally, however, an electron will 'stray' over into a neighbouring metal ion. Thus, at any given instant, a few neutral atoms of fluorine will exist in the lattice. Now, the process of X-ray excitation (in this case K excitation) by an impinging electron is so swift that there will not be time for a redistribution of the valence electrons to occur, thus either a fluorine ion or a fluorine atom may lose a K-electron. If the X-ray excited state of the neutral fluorine atom is non-atomic, the intermediate state will, in spite of this, always be the same For the valence-electron levels will not be affected by the removal of a K-electron from either type of fluorine centre, and the electrons will distribute themselves around the K excited centre so as to give the minimum energy But, if the excited state produced by the removal of a K-electron from the neutral fluorine atom is atomic, then its valence levels belong to itself alone. They will lie well below the levels of the general band in the lattice, and will form a set of sharply defined energy. An electron from the rest of the material can only pass into them, thus altering the K excited state, by getting rid of energy by radiation or otherwise, a process requiring a time at least as long as that needed for the emission of the K-radiation itself. The X-ray excited state produced by removing a K-electron from a neutral fluorine atom therefore remains, in this case, distinct in energy from that produced from a fluorine negative ion in the crystal. A subsidiary value of the excitation energy, W_x^1, will occur besides the main value, W_x^0, and satellite radiation of higher energy than that of the main band will be emitted. It can be seen that there is the possibility of the existence of such atomic excited states by imagining the crystal to become more and more ionic, until we almost reach the condition of an assembly of free ions in which there is only an occasional jump over of an electron into the positive ion.

The band or line character of the satellite radiation will depend on whether the fluorine centre from which the radiation has been emitted has, *after the emission process*, valence levels of non-atomic or of atomic type. The neutral fluorine atom which was originally K-ionized may end as a fluorine *positive* ion, with a valence-electron missing. Since this form of fluorine centre does not exist in the normal lattice, but has merely been created by the electron bombardment in the X-ray tube, its levels must certainly be atomic. On the other hand, it is possible to imagine that, during the actual emission-process, an electron from the general band of levels is drawn, without loss of energy, into the neighbourhood of the emitting atom, and that the final state is non-atomic (corresponding to a single positive hole in the band of levels of the lattice). The determining factor would seem to be the velocity of interchange of electrons in the crystal The time of passage of a conduction electron from atom to atom in a metal is of the order of 10^{-15} sec., but it is clear that the time in an ionic crystal may be considerably greater. On the other hand, the time of the emission process is probably about 10^{-13} or 10^{-14} sec , being smaller for radiation of short wave-length. Thus the K-radiations of fluorine and oxygen, emitted from fluorides and oxides, lying at the short wave-length end of our range, are favourable cases for the occurrence of line satellites, and we shall see that these actually seem to be found. On the other hand, the metal-ion spectra of the oxides lie in the region of long wave-lengths and show band satellites. The degree of polarity of the substance is also obviously important as a determining factor.

The result of this discussion may be summarized as follows·

(1) *If, owing to statistical fluctuations in the number of electrons associated with a given atomic centre in a crystal, this centre can be excited for X-rays in two different ways, in one of which it has valence levels of atomic type, then satellite emission of class (B) will occur.*

(2) *The satellite may be either a line or a band, according to whether the final state, after the emission of the radiation, has atomic or non-atomic character.*

The picture we have put forward has been built up from actual examples in order to provide the simplest possible foundation for understanding the considerable complexity of the results. It is perhaps crude and speculative in detail, but the existence of satellites and multiple bands seems definitely to imply a lack of uniqueness of the state of a partly polar compound when an electron is removed from an inner shell.

4. HALIDES

The halides, especially those of the alkali metals, are the compounds in which the dissociation into positive and negative ions is most complete, and the proportion of the binding energy due to electrostatic interionic forces is greatest. Nevertheless, we shall show that the effect of the lattice binding on the valence levels is considerable, especially for those salts with a comparatively large electron-density.

We have seen in § 3 that the observed spectra represent the actual $N(E)$ curves for the separated groups of s and p valence levels. Owing to the l selection rule and the fact that different types of spectra lie in our wave-length region for the different halogens, we obtain in some cases data relating to the s levels and in others data relating to the p-levels. In the case of chlorides, the L_2- and L_3-spectra of chlorine both represent the $3s$-levels. For the fluorides, we have the K-spectra of fluorine, which give the band of the $2p$-levels. For bromides and iodides, the strongest available spectra of bromine and iodine are M_{45} and N_{45} respectively. There are also weaker M_{23}- and N_{23}-spectra (Siegbahn and Magnusson 1934*b*), but these are not of much interest from our point of view, since they are liable to be broadened by the Auger effects $M_{45} \rightarrow M_{23}$ or $N_{45} \rightarrow N_{23}$ (see Paper I, § 8).

The L_{23}-spectra of chlorine from several chlorides, taken from Siegbahn and Magnusson (1935), are shown in figure 1, together with a new result of ours for LiCl. Our experimental equipment was not particularly suited for the chlorides, but we hope that this third-order spectrum is adequately resolved. It will be seen that, except for LiCl and CsCl, the spectrum consists of four resolved lines. The lack of resolution in the case of LiCl is no more than would be expected from the greater breadth of the lines, but this does not apply to CsCl, for which two of the lines are almost absent. Leaving this peculiarity over for the moment, the results show that each of the L_2- and L_3-spectra consists of *two* lines, the distance between pairs corresponding to an $(L_2–L_3)$ separation of 1·60 eV. These are marked, in a notation to be explained later, L_3^-, L_2 and L_3^0, L_2^0 Since there is no obvious mechanism for splitting the s levels in a cubic lattice, the short wave-length lines of each pair must be identified as *satellites*. As to the structures on the low-energy sides of the main bands, although we cannot vouch for their complexity of form, there is reason to believe them to be genuine.

Four cases of fluorides, are represented in figure 2, which shows the first-order K-spectra of fluorine. We give actual photometer curves, selected from several as characteristic, and inflexions such as those marked by letters have been verified.* The effect of an occasional spot on the photographic plate, which might be misleading, has been removed in the diagram. It will again be seen that, if the 2*p*-levels form a simple band, then a part of the radiation marked K° must be set aside as a satellite. No low-energy bands of appreciable strength occur.

---

* On account of the small scale of the reproduction, the feeble inflexions on both sides of the chief maxima of the curves for NaF and KF appear very slight, but are, nevertheless, genuine.

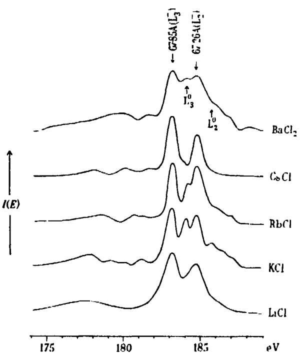


FIGURE 1 Chlorine L_{23} bands from chlorides.

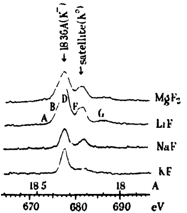


FIGURE 2. Photometer records of fluorine *K*-bands from fluorides.

Turning finally to the case of bromides and iodides, we are faced with a certain amount of spectroscopic complication, since the spectra are of the type $p \rightarrow d$. There are three allowed transitions, namely $p_{\frac{1}{2}} \rightarrow d_{\frac{3}{2}}$ and $p_{\frac{3}{2}} \rightarrow d_{\frac{3}{2}}$ or $d_{\frac{5}{2}}$. Statistically, these have intensities 2 : 4 : 9. First-order photometer curves of the Br M_{45} radiation from several bromides, including AgBr on account of its photographic interest, are shown in figure 3, the N_{45}-radiation of I from CsI is traced in on the same voltage scale. The result for KBr is slightly contaminated by fourth-order L_{23} lines of K (marked K IV). It will be seen that the spectra consist of a number of "lines", more or less overlapping. The case of RbBr has been given by Siegbahn and Magnusson (1934*b*), who have analysed the various transitions involved. We believe that their analysis is erroneous, since it does not postulate a $(p_{\frac{1}{2}} - p_{\frac{3}{2}})$ spin-doublet separation for the levels of the salt which is even approximately equal to the known corresponding separation for the free bromine atom (0·7 eV); we may remove an electron from the Br ion and leave the resulting atom in either p state. The details of our analysis are given in figure 4, in which the M_{45}-spectra of Br and Rb from RbBr are traced on the same volt scale.* It will be seen that the two are very similar and this is also true of the N_{45}-spectra of I and Cs from CsI, in spite of a more pronounced difference in line widths. In the latter case also, the spin-doublet separation is correct, namely 1·3 eV, as for the iodine atom. Owing to the decomposition of the material of the target (§ 2), the Rb and Cs spectra come almost entirely from these elements in metallic form. But this makes little difference, since we are concerned with transitions, not of the valence electrons, but of $4p$ or $5p$-electrons which form a complete shell inside the valence-shell of the Rb or Cs atoms. Br and Rb, I and Cs, are, in the periodic table, only separated by rare-gas elements. Hence the structures of Br^- and Rb^+ should be very similar, differing only in absolute energy, the $4p$-levels being only weakly bound in the case of Br^- but bound with approximately 10 eV in the case of Rb^+. The close analogy between the two spectra shows directly that, since the $4p$-levels of Rb^+ certainly form a closed atomic group, the same applies to the $4p$-levels of Br^- in the solid.

It will be noticed in figure 4 that, both for the Br and the Rb spectra, quite strong features (those marked M^0_{45}) are omitted from our line analysis. These must be put down as satellite which are considerably stronger in the case of Br than that of Rb. Thus the K-spectra of F in the fluorides, the

---

* The inflexion corresponding to the transition $p_{\frac{3}{2}} \rightarrow d_{\frac{3}{2}}$ is weak, but we believe definite, for KBr and RbBr; for CsI, owing to the larger difference between the p- and d-doublets, it is quite evident.

L_{23}-spectra of Cl in the chlorides and M_{45}- and N_{45}-spectra of Br and I in the bromides and iodides all show *satellites which are exceptionally strong.* The satellite of the Rb spectrum in figure 4 may almost certainly be

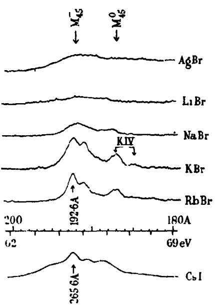


FIGURE 3. Photometer records of bromine M_{45}- and iodine N_{45}-bands from halides.

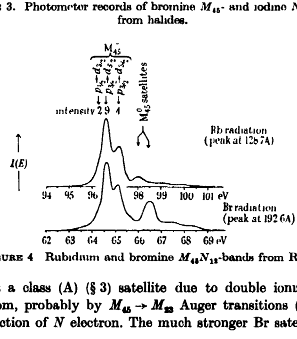


FIGURE 4 Rubidium and bromine $M_{45}N_{12}$-bands from RbBr.

described as a class (A) (§ 3) satellite due to double ionization of the emitting atom, probably by $M_{45} \to M_{23}$ Auger transitions (Paper I, § 8) with the ejection of N electron. The much stronger Br satellite, like the corresponding satellites of the fluorides, must be of class (B), that is an intrinsic property of the salt. Nevertheless, the analogy between the Br and Rb satellites shows that the systems from which they are emitted are similar; consisting namely of complete shells with one M_{45}- and one N-electron missing. This gives support to our view that the Br satellites are to be regarded as emitted from *neutral* Br atoms in the crystal from which an M_{45}-electron has been removed, and hence that these neutral atoms have a real existence in the salt. The satellites of the fluorides and chlorides must then be assumed to have a similar origin. Thus we may describe the main bands of figures 1–3 as 'negative bands', since they correspond to the X-ray excitation of negative ions, while the satellites may be called 'neutral' radiation. This is the reason for our symbols K^-, K^0, L_3^-, L_3^0, etc., marked on the diagrams.

A certain amount of confirmation for this view may perhaps be obtained by applying the free atomic formula (5). This gives the values 2·9 and 2·1 eV for the energy separations of lines emitted after X-ray excitation of negative ions and neutral atoms of fluorine and bromine respectively. The observed separations in the salts are 3·7* and 1·9 eV. The corresponding calculation for chlorides cannot be done, since the energies of the 3s electrons are not known.

The satellite bands of the fluorides are considerably narrower than the main bands. Comparison with the K-spectrum of oxygen from CaO (§ 5) leads one to assume that they may well consist of a pair of rather sharp lines, in fact, some traces of doublet structure can probably be seen in the observed curves. The lack of resolution is very likely experimental. It seems fairly sure that we have here a case of the line satellites discussed in § 3. The doublet character would then be explained as due to the possibility of singlet and triplet K-spectra for an atom such as neutral fluorine. When a K-electron is removed, we have two positive 'holes', both in the intermediate excited and final states, and the interaction between these will give rise to spectroscopic term-multiplicity. The observed separation of 1·2 eV would seem to be of the right order of magnitude. For the bromides, the satellite may also be taken as a line satellite; there is no definite indication of term multiplicity here, but this might well be lost among the complexity of the $N_{23} \rightarrow M_{45}$ transitions For the chlorides, the main spectra as well as the satellites may probably be regarded as consisting of lines, since the 3s-levels are anyhow of more or less atomic character. The structure seen at the high-energy end of the bands of KCl and RbCl is possibly a complication due to term-multiplicity.

---

* For the oxides, the calculation gives 5 eV, instead of 2 eV. We do not know what reliance can be placed on the ionization potential of fluorine, which is involved in both cases.

It will be noticed that the strength of the line satellite depends markedly on the metal ion in the crystal. Thus the satellite is weak for KF, and almost non-existent for CsCl. On our view of § 3, the strength of the satellite radiation depends on the number of negative ions from which, at a given instant, an electron has 'strayed' away to a neighbouring positive ion. This will obviously vary with the character of the positive ion, and in general will be least for the halides which have the most markedly polar structure.

If we wish to obtain the $I(E)$ curve for the main band, in order to determine according to equation (2) the $N(E)$ curve for the band of levels involved, the line satellites must be eliminated. This is easily done, but we also have to consider the projections of the observed curve which occur at energies higher than that of the satellite The question is whether these also are to be rejected as satellites of some sort, or are to be considered as parts of the main band We think that the second alternative is the correct one, this is largely based on the analogous case of the oxides and the reasons will appear in the next section of the paper. On this assumption, the $N(E)$ curves for the p-bands of valence-electron levels in various fluorides and bromides have been drawn in figure 5, together with curves for the 3s-bands in some chlorides which, in contrast to the others, appear simply to have the form of broadened lines The p-bands have total widths which may seem surprisingly great. In fact, we have been very unwilling to recognize the existence of such wide bands and have only been forced to admit them because every alternative hypothesis seems unlikely. The band widths are, however, considerably less than the values found for certain solid elements which have comparable valence-electron densities and the main parts of the $N_p(E)$ curves are confined to a comparatively small energy range, the total width being made up by weak projections of the bands towards high energies. Certain structure effects, marked by letters on figure 5, are evident in the case of the fluorides, in the case of the bromides, we have retained, in figure 5, the $(4p_{\frac{3}{2}} - 4p_{\frac{1}{2}})$ spin-doublet separation, since this has a real existence in the solid. For the 3s-bands of the chlorides, no sign of structure has been indicated in the bands, although some might possibly exist for KCl at the high-energy end of the band.

The explanation of the weak *low-energy* bands of the L_{23}-spectra of Cl from chlorides (figure 1) must be left open. But, since similar structures seem to exist in the L_{23}-spectra of sulphur from the sulphides (§ 6), we may suggest, as a possibility, bearing in mind the fact that $s \rightarrow p$ transitions may have a low absolute probability, that they are due to double transitions. One transition would be the 'forbidden' $3p \rightarrow 2p$, while the other would consist of the formation of an 'exciton' (Mott and Gurney 1940). This would leave the system, after the emission of radiation, in an excited state corresponding to an alternative value of $W(E)$, (equation

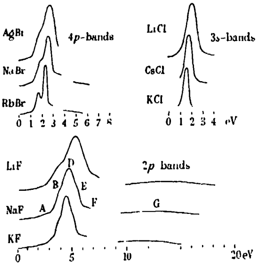


FIGURE 5 Bands of levels for halides.

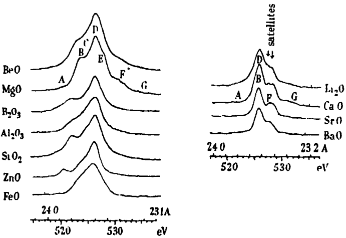


FIGURE 6 Photometer records of K-bands of oxygen from various oxides.

(1)) and would appear to give an energy of the right order of magnitude. The low-energy bands of the L_{23}-spectra of Si from SiO_2 (figure 8) and of sulphur from sulphides (figure 7) may perhaps have a similar origin.

5. OXIDES

For a number of oxides, we have been able to obtain the soft X-ray spectrum of the metal-ion, as well as the K-spectrum of oxygen. For the present, however, we shall leave the metal ion out of account. The oxides, in general, must be considered to be less polar compounds than the halides; the O^{--} ion, in the free state, is unstable, and it is doubtful whether we can suppose that it has any real existence in the lattice. As we shall see, oxides vary considerably in polarity, but, from the spectroscopic point of view, the negative charge on any oxygen centre does not seem to exceed one unit. The oxides are only partly polar compounds, the homopolar component of their binding being often strong.

A number of first-order K-bands of oxygen from various oxides are shown in figure 6,* some remarks on the construction of this diagram have already been made in § 2 They fall into two groups, in the second group the curves show a more or less resolved line doublet on the high-energy side of the peaks. The fact that MgO and CaO have the same crystal structure shows that this is not the distinguishing factor between the two groups. It is therefore very difficult to account for the difference between the two groups except on the assumption that the line doublet is a satellite, such as we have postulated for the halides. And indeed the form of the K-bands of oxygen from these oxides is very similar to that obtained for the K-bands of fluorine from the fluorides. Compare the features marked by letters on the curve for CaO with those marked for LiF on figure 2.

We shall therefore proceed on the assumption that the observed doublet is a line satellite, the doubling being due to the multiplet splitting of the K-spectrum, in this case probably into the doublet and a quartet spectrum corresponding to a complete shell with three missing electrons The distinction between the two groups of oxides is mainly due to the lower electron density in the latter group which makes them more polar compounds, like the halides (see § 8). We must assume that, for the first group of oxides, in spite of the fact that the number of valence electrons associated with a particular oxygen centre is not a constant, no intermediate K excited state of atomic character (§ 3) can exist, while for the second group, such atomic K excited states do occur. They are probably to be regarded, as we shall see in § 7, as due to the possibility of the momentary existence of *neutral* atoms of oxygen in the crystal, owing to the straying of electrons away from the negative ions. The total intensity of the satellite lines is, in the case of CaO, about 5 % of the intensity of the main bands. We may therefore assume that at any moment, at least 5 % of the oxygen centres are atoms. We only obtain a minimum estimate because of the possibility that transitions of valence electrons may destroy the atomic character of the K excited state (§ 3).*

---

* See also the recent paper of Tyrén (1939) which gives many of these results Indeed, his better resolving power shows itself especially in the sharpness of the double peaks on the high-energy side of the maxima for CaO, SrO, BaO, which we have resolved only for CaO. Tyrén interprets the peaks of the curves as a set of X-ray lines, on the analogy of hard X-rays. It is obvious that this is inconsistent with the standpoint of the present paper. We found no sign of a diffuse subsidiary band at 24,44 A, reported by Magnusson (1938) and others. If genuine, it must have a very low intensity.

It might be thought that certain features of the spectra of the first group of oxides ought to be regarded as satellites. The line satellites of the halides and those of the second group of oxides have been identified by their common characteristics; the energy separation from the main band is constant in a group and the form of the curves similar. It is seen that there are no such common features among the curves of the oxides of the first group, and thus there are no line satellites. But the further question arises of whether the weak high-energy parts of the curves of both groups of oxides should be taken as band satellites, of which we shall give examples in § 7. There is no obvious resemblance, but this might be due to their faintness in the present case. A band satellite may be expected to be considerably further from the main peak than a line-satellite, in fact, according to equation (4), the separation would be about 15 eV. This probably cannot be trusted. But, to take a particular case, the feature G of the curve of MgO (figure 6) is separated from the main peak by a different energy from that of the apparently corresponding feature G of CaO. The feature F of MgO has no counterparts on the curves grouped below it. Also F, and in fact all the features of the K-spectrum of oxygen from MgO, appear to be reproduced, as we shall see in § 8, in the L_{23}-spectrum of Mg emitted from the same compound. Thus, although this evidence cannot perhaps be regarded as conclusive, it seems that *no part of the observed K-spectrum of oxygen from* MgO *is to be identified as a satellite*. If this is granted, it appears to follow by analogy that there are no band satellites in the curves of the oxides of either group, and also in those of the fluorides (figure 2). This is the evidence on which the long high-energy projections of the halide bands have been included in figure 5. Also, on this basis, the curves of figure 6, with the line-satellites eliminated from those of the second group as indicated roughly by the dotted lines, represent directly the $N(E)$ curves for the $2p$-bands of levels of the various oxides, according to equation (2). We shall, in § 7, describe results which are interpreted as giving the $2s$-bands of some of these oxides, and a discussion of the various band forms will follow in § 8.

---

* However, the L_{23}-spectrum of Ca in CaO, obtained subsequently to the work described in this paper, can probably be interpreted as showing that the separation into ions, singly charged, is nearly complete.

6. SULPHIDES

Since sulphur is the homologue of oxygen in the next group of the periodic table, it is convenient now to consider the case of the sulphides, which resemble the oxides. The advantage of this is that for the oxides we have only the K-spectra and these give information only about the p-levels. On the other hand, for the sulphides, we have L_{23}-spectra, giving information regarding s- and d-levels (equation (2)). Unfortunately, the K-spectra of the sulphides lie in the region of short wave-lengths (about 5 A) and have not been observed.

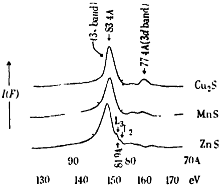


FIGURE 7. Sulphur L_{23} radiation from sulphides.

Photometer curves of three examples of sulphides are shown in figure 7. Since they are taken from spectra of first and second order, they have been traced on the same volt scale. The most striking characteristic is that, unlike the oxides, the sulphides have bands consisting of *two parts*. The high-energy part is weak and shows signs of structure, while the low-energy part is strong and relatively formless. The separation between the two parts is about 11 eV, and the band-width of each is of the order of 10 eV. The $(L_2 - L_3)$ separation for sulphur may be noticed in a sharp fall of the intensity at the high-energy end of the stronger band of ZnS, which shows clearly double. The other sulphides do not show this peculiarity.

On account of the selection-rule, it seems certain that we must identify the main band of the L_{23} (or $2p$) spectra as representing the $3s$-band of the sulphur ions. The high-energy band must probably be interpreted, not as a satellite band, but as representing a few levels with wave functions most likely of $3d$-type, which are to be found in the band of levels whose wave functions are mainly of $3p$-type and so do not show in the spectrum. This band therefore corresponds to the band of the K-spectrum of oxygen from oxides; there is nothing in the latter corresponding to the main sulphide band, since $2s \rightarrow 1s$ transitions are forbidden. Though it is weak, and so perhaps some details may have been missed, the high-energy band of the sulphide spectra show the same kind of intricate structure found in the K-spectra of the oxides. The main band is, on the other hand, except in the case of ZnS, completely formless; and even for ZnS, it is formless except at its high-energy end. The reason for this peculiarity is rather obscure: it may be that a few $3d$-levels are mixed in the $3s$-band and that the sharp features of the main band represent these. This would mean that the interaction is greater for ZnS, and makes its spectrum tend to resemble that obtained from the element sulphur, for which all the levels are more or less admixed and cover the range of both sulphide bands (see Paper I, § 5).

The most obvious feature of the $3s$-bands of the sulphides is their shapelessness. This has to be explained, since there seems to be no reason why the effect of the lattice should not show itself in a characteristic structure of the $3s$-band, just as for the $3p$-band. We therefore need a mechanism which both smears out the structure of the $3s$-band and broadens it. Fortunately, a similar smearing-out of the features at the low-energy end of the observed bands of metals and element-insulators has been described (Paper I, § 9). The mechanism used to explain it is an Auger effect in which an electron in a level near the top of a band of levels may make a transition into a vacancy near the bottom of the band. Owing to this possibility, the state which corresponds to a vacancy near the bottom is broadened, and the resulting width at half-maximum is shown to be of the order of 1–2 eV. In Paper I, § 8, it was also shown that the $2s$-levels of Mg and Al (which are 'inner atomic' levels, not valence-levels) are broadened to the extent of about 2 eV by the possibility of Auger transitions from the $2p$-levels into them, with the ejection of conduction electrons. The case of the $3s$-

levels of sulphur in sulphides (and also that of the 2*s*-levels of oxygen in oxides) lies intermediate between these two examples of level-broadening, and we have every reason to expect a similar effect. The broadening can only be guessed, but certainly a level width of the order of 2 eV may be assumed. This would have the required effect of smearing out any characteristic structure of the bands, and would account (as in Paper I) for the 'tailing' of the bands towards both low and high energies. A has been emphasized in Paper I, this broadening of the levels must not be considered as a spurious effect. The Auger effect is one of the important causes of the spreading-out of sharp atomic levels into a band in the solid. But it is an effect not taken into account in theories of the solid state, and so has to be separated out if one is thinking only in terms of the interaction between neighbouring atoms In contrast to these *s* bands, the 2*p*-bands of the oxides and fluorides are remarkably free from any effects of level broadening This is probably due to the fact that the only possible Auger transitions within these bands would be of the type $p \rightarrow p$, and these may well be very improbable.

In the following section of the paper, we shall show that the assumption that the 2*s*-band of levels in an oxide has properties similar to those of the 3*s*-band of a sulphide is supported by strong indirect evidence.

7. METAL-ION SPECTRA FROM OXIDES, ETC.

So far, we have only made a bare reference to the metal ion K- and L_{23}-spectra which are available for BeO, B_2O_3, MgO, Al_2O_3 and SiO_2. We also have spectra of both components in the cases of BN and SiC For most of these compounds, the spectra show a high degree of complexity, thus the K-spectra of the metals in BeO, B_2O_3 and Al_2O_3* consist of three very well-marked distinct bands, the weak low-energy band being probably a duplicate of the main band; while the L_{23}-spectrum of Mg in MgO presents an obvious case of band duplication.

The curves are shown in figure 8, and the K-spectra of the relevant negative ions have been traced in, in an arbitrary position on the voltage scale. The relative positions of the corresponding bands from the pure metals are indicated by the emission edges or 'false edges' (Paper I, § 6). In the cases of MgO and Al_2O_3, there seems to be an identity of the characteristics of the K-band of oxygen with each separate part of the L_{23}-spectra of the metal ions; this is most evident for MgO, and in spite of the fact that we were obliged, owing to technical difficulties to rely on rather weak spectra, the apparent identity was carefully checked using several

---

* The detail observed at the high-energy end of the other K-bands is missing in this case. The radiation investigated by Karlsson and Siegbahn (1934) lies at a short wave-length (8 A) and it is possible that the experimental resolution was not adequate.

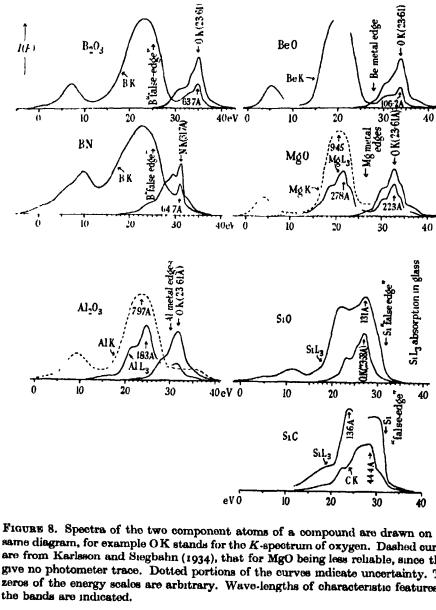


FIGURE 8. Spectra of the two component atoms of a compound are drawn on the same diagram, for example O K stands for the K-spectrum of oxygen. Dashed curves are from Karlsson and Siegbahn (1934), that for MgO being less reliable, since they give no photometer trace. Dotted portions of the curves indicate uncertainty. The zeros of the energy scales are arbitrary. Wave-lengths of characteristic features of the bands are indicated.

exposures. In the cases of BeO, B_2O_3 and BN,* a sort of correspondence between the K-spectra of the negative ions and the high-energy part of the K-spectra of the metal ions is established. In the remaining cases, there is no very obvious correspondence.

Let us first consider the case of a typically homopolar compound SiC. We have the K-spectrum of C and the L_{23}-spectrum of Si. If the compound is strictly non-polar, with one band of valence levels held in common by all the lattice points, it is clear that one can expect no more resemblance of form between these spectra than is found in the K- and L_{23}-spectra from the respective elements, diamond and Si, which have the same crystal structure. In Paper I, § 10, it was shown that there is no very obvious similarity between these spectra, a fact interpreted, according to § 3, by saying that the K-spectra give $N_p(E)$ and the L_{23}-spectra $N_{s+d}(E)$, and that the overlapping systems of $(s+d)$ and p-levels have quite different characteristics. So, in SiC, we cannot expect any particular similarity between the K curve of carbon and the L_{23} curve of silicon, apart from an identity of band width. If the K-spectrum of Si were available, it would no doubt show features nearly identical with those of the K-spectrum of C.

The remaining cases of figure 8 are examples of semi-polar compounds, though the characteristics of the bands of SiO_2 show a certain resemblance to those of the homopolar SiC. We shall return to this exceptional oxide subsequently; for the present we shall deal with the other cases which all show similar features. Of all the properties of soft X-ray bands dealt with in Paper I or in this paper, those of the metal-ion spectra of these semi-polar compounds are the most complex. Nevertheless, as stated above, we have several general conclusions to draw upon, and it seems worth while to put forward the following suggestions.

From what has been said in §§ 5 and 6, we may take the system of valence-electron levels in the oxide as known, at least provisionally. It consists of a band of $2p(\mathrm{O})$ levels, of the order of 13–20 eV in width with an intricate $N(E)$ curve; and, only slightly overlapping it, a band of $2s(\mathrm{O})$ levels about 10 eV in width and of a formless structure. Ignoring for the moment the observed band duplication, the similar forms of the Mg L_{23}-spectrum and the K-spectrum of oxygen in MgO show conclusively that in this case *we have to deal with the effect of symmetry-interchange of wave-functions*, already referred to in § 3. Those levels of the lattice which have $2p$ character when referred to the oxygen centres, have $3s$ character referred to the $\dot{Mg}$ centres; and though less distinctly seen, the L_{23}-spectrum of Al in Al_2O_3 shows similar characteristics. On the other hand, the K-spectra of Mg and Al in their oxides, as well as the similar, but more accurately observable K-spectra of Be and B in BeO and B_2O_3 show different peculiarities. The main feature of the band (whose duplication at lower energies will, for the moment, be ignored) is a formless structure, almost certainly to be interpreted as the $2s$(O) band which, owing to the selection rule, does not show in the K-spectrum of oxygen. We have therefore another case of the symmetry-interchange effect, the $2s$(O) wave-functions becoming, for example, $2p$(Be) wave functions. But this is not all, at higher energies in the observed K-spectra, we have weak structures which show *some* of the properties of the oxygen K-spectra. Hence we must assume that some of the $2p$(O) wave functions correspond to $2p$(Be) wave functions, the probability of such a correspondence being however small (of the order of 10 %). Also, in order to explain the distortion of the oxygen K-spectrum in the metal ion K-spectra, one must postulate that this probability is a function in general decreasing with the energy E of the level concerned, but perhaps also depending on the character of the wave function giving rise to the level.

---

* It was not found possible to obtain spectra of BN free from B_2O_3, and the bump on the curve at 33 eV is almost certainly due to the latter compound.

The substance SiO_2 shows different, and in some respects opposite properties. The L_{23}-spectrum of Si shows little similarity with the relevant K-spectrum of oxygen, though perhaps some correspondence can be traced at the high-energy ends of the bands. In any case, there are many $3s$(Si) levels which do not correspond to $2p$(O) levels and therefore presumably must correspond to $2s$(O) levels. The L_{23}-spectrum of Si in SiO_2 may probably be compared with the K-spectra if the metal ions in the other oxides dealt with, except that there is more overlapping between the parts of the band which appear to correspond to $2s$(O) and $2p$(O) levels. Further, a private communication of Dr Farineau indicates that the K-spectrum of Si in SiO_2 is very similar to, or identical with, the oxygen K-spectrum. We may therefore conclude that in SiO_2, the symmetry character of a wave function of given energy near the Si nuclei is mainly the same as it is near the O nuclei.*

If our analysis is correct, we may conclude that the main bands of the K-spectra of the metal ion in BeO, B_2O_3, MgO, represent the bands of $2s$(O) levels. Like the $3s$-bands of the sulphides, these are wide and diffuse;

---

* Subsequent experiments indicate that CaO and TiO_2 are rather similar to SiO_2 in this respect. On the other hand, unlike SiO_2, they show high-energy satellite bands, as do also the oxides of the Fe-group, Cr_2O_3, MnO_2 and Fe_2O_3 (see Gwinner (1938) for the latter).

they are separated from the mean energies of the $2p$-bands by 11 or 12 eV. The values for Al_2O_3 and for the corresponding separation in BN are smaller, namely, about 7 or 8 eV and the separation of the $2s$(O) and $2p$(O) bands in SiO_2 is probably only about 6 eV. Though no accurate data exist, the separation of $2s$- and $2p$-levels in the free oxygen atom is thought to be about 8 or 10 eV

We must finally consider the *band-multiplicity* which is shown by all the substances represented in figure 8, with the exception of SiC. It is difficult to resist the conclusion that we are dealing with cases of band satellites, in the sense of § 3, in which the intermediate X-ray excited state is duplicated, but the final state is always the same and is formed simply by the removal of a valence electron from the general band of levels in the crystal. The great strength of these satellites shows that they must undoubtedly be of class (B). It may be noted that none of the substances represented in figure 8 have any satellites of their oxygen, etc., K-spectra. This means that secondary intermediate states of atomic character can often exist for the metal centres, without corresponding intermediate states being possible for the oxygen centres.

In the cases of MgO and Al_2O_3 the main bands of the L_{23}-spectra of the metal ions lie in approximately the same position in the spectrum as the L_{23}-bands from the corresponding elements. Since the $2p$(O) levels are bound with about the same energy as the conduction electrons in the metals, this means that the structure of the intermediate state is similar in the metal and in the chemical compound. Hence, in a rough sense, we may probably conclude that the emission of these main bands is from centres which, before the removal of an L-electron, were originally 'neutral atoms', not 'positive ions'. On the other hand, the satellites of these bands will correspond to emission from positive ions. This reasoning cannot be applied to the K-bands of BeO, B_2O_3, BN, MgO and Al_2O_3, because here we are dealing with transitions from the $2s$(O) levels, which lie comparatively low. Hence one may expect a shift of the bands which correspond to emission from neutral atoms of the order of 10 eV towards lower energies as compared with the K-bands of the respective elements. The weaker low-energy bands are therefore, for these substances, those which correspond to emission from neutral atoms, while the bands which lie in about the same position as those emitted from elements actually correspond to emission from positive ions.

In this sense, we may say that oxides like BeO and MgO are substances composed of a mixture of neutral atoms and singly-positive and negative ions. The number of electrons associated with a Mg or Be centre is probably either one or two, and we have a statistical distribution of centres of these kinds. This appears to indicate a strong non-polar constituent in the binding of these oxides.

There would seem at first sight to be two difficulties in this point of view: the first is that the energy separation of the main and satellite bands in the case of MgO is different for K- and L_{23}-spectra. This is probably due to the fact that the screening of the nuclear charge by a K- and an L-electron are not the same, and therefore the two values of W_x (equation (1)) differ by less for an L_{23}-spectrum than they do for a K-spectrum. The second difficulty is that the ratio of the intensities of the two bands which correspond to emission from neutral atoms and positive ions is not the same for a K- and an L_{23}-spectrum, the neutral band being relatively much stronger for the latter. But it must be remembered that the emission of radiation is not the only process by which an atom can get rid of the energy of excitation,* and thus the relative intensities of the observed bands do not necessarily represent the relative numbers of the two kinds of X-ray excited atoms. Thus we can only conclude that the mean number of electrons associated with a Mg centre in MgO lies between about 1/3 and 3/4 and is probably of the order of 1/2. The other oxides represented in figure 8 seem to be similar, again with the exception of SiO_2. The L_{23}-spectrum of silicon from this substance occupies almost exactly the same spectrum range as that from the pure element; there is no high-energy band at all. This does not necessarily mean that quartz is non-polar. It means simply that only one X-ray excited state exists, and this may be interpreted according to § 3 in saying that no intermediate state of atomic character can occur. Nevertheless, as already emphasized, the band of quartz is very similar to that of a covalent compound. We have previously suggested that the weak band of low energies, which might be taken as due to a few negative ions of Si, may really have an origin similar to that of the low-energy bands in the L_{23}-spectra of the chlorides.

We may summarize this intricate discussion of the evidence provided by the metal-ion spectra in stating that they appear to show·

(1) That the $2s$(O) bands of the oxides appear to be, like the $3s$-bands of the sulphides, diffuse and fairly wide. Except in the case of SiO_2, they are separated from the $2p$(O) bands by an energy interval of the order of 10 eV.

---

* For example, radiation from the neutral K excited state Mg ($1s.2s^2.2p^6.3s^2$) might be assumed relatively improbable owing to the selection rule, the excess energy being usually taken away by an Augor process with the ejection of a $3s$-electron.

(2) The compounds show the symmetry-interchange effect (namely that a wave-function which has *s*-like symmetry near one type of centre in the lattice may have symmetry of *p*-type near the other type of centre) in more or less degree. The effect is small in the case of SiO_2, considerably greater for BN and Al_2O_3, and greater still for the remaining oxides BeO, MgO, B_2O_3.

(3) The duplicate character of the spectra is explained by assuming that the atoms which emit the radiation are not all the same. In conjunction with the results of §§ 4 and 5, this leads to the general conception that a semi-polar compound is to be regarded as an assembly of ions and atoms, the charge on any centre being liable to change in a time of the order of 10^{-13} sec. The halides consist mainly of negative halogen ions and positive ions with an admixture of a few per cent of neutral atoms, which represent the covalent constituent of the binding. The oxides consist of singly negative ions and corresponding positive ions, but there seems usually to be a fairly strong admixture of neutral atoms, except in the case of the oxides of lowest density, like CaO, SrO and BaO, which appear to be built up very similarly to the fluorides, with a polarity corresponding to nearly one electron-charge. The fact that they have exceptionally high refractive indices in relation to their densities is perhaps connected with their greater degree of polarity. At the other end of the scale, the oxide SiO_2 has spectra similar to those of a covalent compound like SiC.

8. GENERAL PROPERTIES OF THE BANDS

We shall now attempt to draw a few conclusions from the forms of the observed bands. In tables 1–3, we give the band-widths of the *p*- and *s*-bands of the various compounds investigated. Table 1 gives the values for the 2*p*-bands of the fluorides and for the 4*p*-bands of the bromides, table 2 those for the 2*p*-bands of the oxides and table 3 those for the 2*s*- and 3*s*-bands of the oxides, sulphides and chlorides. The somewhat arbitrary end-points chosen have been, where possible, indicated by dots on the photometer-curves (figures 2, 3, 6 and 7). The points at the high-energy ends of the *p*-bands depend, of course, on our assumption of § 5 that the weak observed projections are not satellites. If this is wrong, the band widths for the fluorides would have only about half, and the other compounds upwards from two-thirds of the values listed. The methods of defining the end-points at the low-energy ends have been discussed in Paper I, § 10, where it has been pointed out that, when there is 'tailing' of the bands, an arbitrary definition is needed. The tailing of the *p*-bands is certainly small, and so we have measured practically to the extreme limits of the curves. But, as we saw in § 6, it is large for the *s* bands, and we have extrapolated their linear portions to zero to get the low-energy end-points. The high-energy ends of the *s*-bands (except for that of ZnS) seem also subject to the tailing, and the widths given represent values in which this effect has been largely eliminated. We give also, in the tables, values of the

TABLE 1. 2*p*-BANDS OF FLUORIDES AND 4*p*-BANDS OF BROMIDES, ETC.

Compound	LiF	NaF	KF	MgF_2	LiBr	NaBr	KBr	RbBr	AgBr	CsI
Line-width, eV	2·7	2·0	1·7	3·1	1·2	0·75	0·55	0·45	1·1	0·7
Band-width, eV	18	16	15	19	>7	6·5	6·1	5·6	>7	6·5
Sommerfeld band-width	21	17	14	20	11	9	8	8	10	7
ρ	7·0	4·8	3·6	6·3	2·5	2·0	1·6	1·5	2·3	1·3

TABLE 2. 2*p*-BANDS OF OXIDES, ETC.

Compound	Li_2O	BeO	B_2O_3	MgO	Al_2O_3	SiO_2	CaO	FeO	ZnO	SrO	BaO	BN
Line-width, eV	2·7	5·5	3·5	5·5	4·0	4·0	2·2	4·3	3·0	2·0	2·2	5·0
Band-width, eV	16	23	19	21	21	20	15	19	18	14	14	20
Sommerfeld band-width	17	22	17	19	20	18	16	18	17	14	13	18
ρ	4·8	7·4	4·8	6·0	6·6	5·3	4·4	5·3	4·7	3·7	3·3	5·7

TABLE 3. 2*s*-BANDS OF OXIDES, ETC., 3*s*-BANDS OF SULPHIDES AND CHLORIDES

Compound	BeO	B_2O_3	MgO	Al_2O_3	BN	MnS	Cu_2S	ZnS	LiCl	KCl	RbCl	CsCl	$BaCl_2$
Line-width, eV	8	9	7	8	10	3·7	3·2	3·8	1·3	0·65	0·65	0·75	1·0
Band-width, eV	12	13	10	13	14	7·5	7·0	8·0	—	—	—	—	—
Sommerfeld band-width	11	8·5	9·5	10	9·0	5·8	5·0	5·4	5·8	4·2	4·1	4·4	5·1
ρ	7·4	4·8	6·0	6·6	5·7	3·0	2·4	2·7	3·0	1·9	1·8	2·0	2·5

'line widths', namely the widths of the bands at half of the maximum intensity, as is usual in measuring broadened spectrum lines, in the case of bromides, these refer to one line of the *p*-doublet. Especially for the *s*-bands, these can be more accurately determined than the full band widths. The quantity ρ represents the density of oxygen, halogen, or sulphur atoms in the crystal, it has usually been calculated from the lattice-constant, the volume of the spheres occupied by the positive ions as determined by crystallography being subtracted from the total volume. In the cases of the more complex oxides and of the sulphides, ρ has been calculated from the actual density, and the small correction for the size of the doubly positive ion (which anyhow is of doubtful significance) has merely been estimated. In this way, by allowing 6 electrons per negative ion in a p-band and 2 electrons per negative ion in an s-band, we obtain at least a rough value for the effective electron-density. So we are able to calculate the Sommerfeld band-width according to the crude 'electron-gas' theory of solids. This quantity, though certainly not strictly applicable in the case of insulator-bands, yet provides a standard against which the measured band-widths may be compared. We have shown in Paper I, § 10, that, in the case of simple crystal structures, the Sommerfeld values are never very far from the real band widths of the element insulators, but are always somewhat smaller. It will be seen that it gives a close approximation to the $2p$-band-widths of most of the compounds, thus showing generally the effect of the electron density. In fact, the results show that valence electrons must move in a semi-polar compound with velocities almost as great as they do in a metal. They probably move from negative ion to negative ion, the positive ions being more or less insulated, but the very low intensities at high energies of the bands of the more ionic compounds show that the number of electrons with large energies is very small. The results for the s-bands are not accurate, but those for the chlorides show that these bands too are sensitive to the electron density. It is rather surprising that the $2s$-bands of the oxides and the $3s$-bands of the sulphides are wider than calculated; it might be thought that these bands of levels, which we have shown to lie about 10 eV below the corresponding p-bands, would be narrower. But in fact, the tables suggest that the interaction of the $2s$-electrons in neighbouring atoms is even greater than that of the $2p$-electrons.

The $2p$-bands of the oxides and fluorides show different forms. One would expect these to be related to the crystal structure. After an initial rise as $E^{\frac{1}{2}}$ at the low-energy end (Paper I, § 9), the bands may be expected to be built up of a series of 'peaks', corresponding to characteristic energies of reflexion of electron-waves by sets of planes in the lattice (see Paper I, §§ 3 and 10). The complicated internal structure of the Brillouin zone which contains eight electrons implies that there will be a number of overlapping peaks in the curves for the halides and oxides, as indeed is observed. Glancing at the curves for oxides (figure 6) one would be tempted to distinguish B_2O_3, Al_2O_3, SiO_2 and ZnO from the remainder on account of the shape of their curves. These all have, except for B_2O_3 whose structure is unknown, various forms of hexagonal structure, while the remainder, except for BeO, are cubic. The case of BeO raises a difficulty, since its structure is given as close-packed hexagonal, like ZnO. Since the energy difference of a cubic allotropic modification (like MgO) must certainly be small, we think it is fairly safe to assume that our finely ground powder was either cubic, or was rendered so by the electron bombardment in the X-ray tube, perhaps simply by becoming hot.

From the fluorides, the tetragonal MgF_2 can perhaps be picked out by a difference in the number of the inflexions (figure 2), although this is not certain. The remainder have face-centred cubic structure, and thus are like MgO, CaO, SrO, BaO, and, with the above assumption, BeO. We have emphasized the similarity of their $2p$-bands by the lettering on figures 2, 5 and 6. But they are not all identical in the intensity of their inflexions. Thus MgO, NaF, and KF have a strong 'shoulder' of each side of the peak (features marked B and E) while BeO and LiF have one only, B; and, in the cases of CaO, SrO and BaO, the shoulders are hardly visible at all. These differences are not due to experimental error, and so must represent a genuine influence of the nature of the compound. It must be remembered that, as in the case of X-ray diffraction, the intensities of reflexion of electron waves will depend on the nature of the metal ion forming the compound. Thus identical bands cannot necessarily be expected from substances with the same crystal structure. One might consider that the pairs MgO and NaF, BeO and LiF would have bands very similar to one another, on account of the fact that the atoms forming them are neighbours in the periodic table. And indeed, a close examination of the curves shows that the $2p$-bands are remarkably alike, in spite of a considerable difference in scale.

A difference present in these comparisons is, that, in the fluoride bands, a much larger part of the band width is occupied by regions of very low intensity. CaO, SrO, BaO, the oxides with the lowest electron densities, are similar to the fluorides in this respect, though the high-energy projections are not quite so long. The ratio of the band width to the line width (tables 1 and 2) gives a measure of this effect. It is seen that it is about 10 for the bromides, 7 or 8 for the fluorides and for CaO, SrO, BaO, and ranges down to about 4 for oxides such as MgO. This ratio seems to be a measure of the degree of polarity of the compound. If we consider the approximation of building up a crystal out of the constituent ions, we shall obtain a line broadening of the levels dependent on the electron density and on the binding energy of the valence electrons. If, on the other hand, we regard the binding as covalent, we shall obtain a band width approximately given by the Sommerfeld formula. It seems as though both types of approximation have to be taken into account; in the case of a very ionic compound, the first applies to the majority of the levels, and the second only to a very small number. The polar nature of the compound has only a small effect on the total band width, although it has a considerable effect on the line width.

Similar projections of the bands towards high energies have been described in Paper I, § 10 for element insulators, although these do not reach so far out. They have been called 'band-ends', and in some cases a parabolic rise of the curve backwards from high energies can be traced. They correspond to the filling up of the corners of the Brillouin zone. Now, in the face-centred cubic structure, the zone containing 8 electrons has two kinds of corner; hence we may expect two band-ends for the halides and the face-centred oxides. Two such band-ends seem to be present in the curves for all these substances. The scale of the band-ends gives the 'effective mass' (Paper I, § 10) of a 'positive hole' in a position in k-space corresponding to a corner of an otherwise filled zone The effective masses appropriate to the two band-ends of the cubic oxides and halides may be given very roughly as $1/2m$ and $1/5m$ respectively. Similar values for compounds have been postulated by Fröhlich and Mott (1939) from data on conductivity. The hexagonal oxides probably have Brillouin zones too complicated for the band-ends to be resolved, and hence the high-energy ends of their bands take the form of a smooth curve. In this respect they are similar to certain elements with complex crystal structure, for example, boron.

The experiments were carried out in the H H. Wills Physical Laboratory, University of Bristol, during the year 1938, and we are very grateful to Professor A. M. Tyndall, F.R.S., for making our collaboration possible. One of us (O'Bryan) was responsible for about two-thirds of the experimental work, but he need not necessarily be held responsible for the development of all the ideas which sprang from our year of co-operation. We should like to thank Professor N. F. Mott, F.R.S., for much help in discussion, and Dr J. E. Johnston for some assistance with the experiments.

REFERENCES

Broili, H , Glocker, R. and Kiessig, H. 1934 *Z. Phys.* **92**, 27.
Farineau, J. 1938 Thesis (Paris).
Fröhlich, H. and Mott, N. F. 1939 *Proc. Roy. Soc.* **171**, 496.
Gwinner, E. and Kiessig, H. 1937 *Z. Phys.* **107**, 449.
Gwinner, E. 1938 *Z. Phys.* **103**, 523.
Hautot, A. 1934 Thesis (Liége).
Hautot, A. and Serpe, J. 1937 *J. Phys. Radium*, **8**, 175.

Karlsson, H. and Siegbahn, M. 1934 *Z. Phys.* **88**, 76.
Magnusson, T. 1938 *Nova Act. Reg. Soc. Sci. Upsal.* (IV), **2**, no. 3.
Mott, N. F. and Gurney, R. 1940 *Electronic processes in ionic crystals.* Oxford.
Siegbahn, M. and Magnusson, T. 1934*a* *Z. Phys.* **87**, 891.
— — 1934*b* *Z. Phys.* **88**, 559.
— — 1935 *Z. Phys.* **95**, 133.
Skinner, H. W. B. 1932 *Proc. Roy. Soc.* A, **135**, 84.
— 1939 *Reports on Progress in Physics*, **5**, 257. (The Physical Society.)
— 1940 *Trans. Roy. Soc.* A, **239**, 95.
Skinner, H. W. B. and Johnston, J. E. 1937*a* *Proc. Roy. Soc.* A, **161**, 420.
— — 1937*b* *Proc. Camb. Phil. Soc.* **34**, 109.
Tyrén, F. 1937 *Ark. Mat. Astr. Fys.* **25** A, no. 32. Stockholm
— 1939 *Medd. Upsala Univ. fys. Instn.* (Volume number not indicated.)
Yoshida, T. 1936 *Sci. Pap. Inst. Phys. Chem. Res. Tokyo*, no. 621, 243.

The distribution of autelectronic emission from single crystal metal points

I. Tungsten, molybdenum, nickel in the clean state

By M. Benjamin and R. O. Jenkins

Communication from the Staff of the Research Laboratories of The General Electric Company, Limited, Wembley, England

(*Communicated by R. H. Fowler, F.R.S.—Received* 29 *April* 1940)

(Plates 4–9)

E. Muller has shown that the angular distribution of field emission from fine metal points can be related to the crystal structure of the metal. He also suggested that the modifications in the pattern from tungsten which took place when the point temperature was raised could be attributed to a movement of the surface atoms which caused local changes in the work function.

Further experiments are described in this paper using tungsten, molybdenum and nickel. It is shown that this surface mobility occurs above 1170° K for tungsten, 770° K for molybdenum, and 370° K for nickel The changes in emission distribution, which are very much more marked for molybdenum and nickel than for tungsten, can be completely explained in terms of a change in the geometrical shape of the point. This is brought about by the action of the high external field upon the mobile surface atoms. The flash-over phenomenon in high vacuum is discussed and an explanation offered in terms of these observations. The effect of small traces of gas is illustrated, and it is shown that the gas film is removed at the relatively low temperature of 620–670° K.

Several possible suggestions are put forward to explain the dependence of the emission distribution on the crystal structure of the metal, but no definite conclusion can be reached with the data at present available.

1. INTRODUCTION

E. Müller (1937, 1938) has published a description of the cold emission patterns obtained from a fine tungsten point when in the clean state, and when oxygenated or bariated, in which he shows that the angular distribution of emission gives a pattern which is related to the crystallographic structure of the metal. The effect of adsorbed monomolecular films can also be studied in great detail. We have carried out similar work with tungsten, molybdenum and nickel, and have studied the patterns from the clean metal surfaces, and the patterns from surfaces contaminated with barium, sodium and thorium. Preliminary results have been reported (Benjamin and Jenkins 1939). In this paper we deal with the clean metals only, and offer what we believe to be a complete explanation of the changes observed in the emission patterns in terms of the alterations of shape of the emitting point caused by the treatment to which it is subjected.

The dependence of the emission distribution on the crystal structure is also discussed.

2. EXPERIMENTAL PROCEDURE

(a) *The tube*

The essential feature for the removal of electrons from a cold metal is the presence of a very high field at the surface of the metal. In the case of two concentric spheres, one very large compared with the other, maintained at a potential difference V, the field at the surface of the inner sphere is V/r where r is the radius of that sphere. Very high fields can thus be produced with convenient voltages by making the radius of curvature of the point sufficiently small. Preliminary experiments showed that very wide angle beams of electrons could be obtained using a tube of the type shown in figure 1. A spherical bulb B, radius 5 cm., is coated over its lower portion with a graphite film connected to a metal lead A. The upper half of the bulb is sprayed with a thin film of zinc orthosilicate (willemite) W. The fine metal point P, is formed on the end of the filament PQ, of 0·05 mm. diameter wire and 3 mm. in length, and is placed at the centre of the sphere. The filament is welded at Q to the centre of a heater H, made of 0·1 mm. diameter tungsten wire in the form of an inverted V. The electron beam forms a solid angle of about 130°, and the pattern can be seen over the upper half of the bulb.

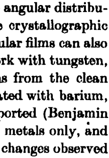


FIGURE 1. Section through experimental tube.

(b) *Operational requirements*

The tube contains a barium getter G, which consists of a strip of tantalum with a V groove filled with a barium-beryllium oxide compound. The strip is arranged so that the getter deposits only in the neck of the bulb. The getter has been described elsewhere (Benjamin and Jenkins 1938). On heating the tantalum, extremely pure barium is produced by reduction of the oxide compound.

The actual weld of the pointed filament was carefully made, so that unwanted sharp points were avoided, and all other welds were covered with glass sleeves to eliminate extraneous cold electron emissions.

(c) *Preparation of the points*

The points were prepared in every case by electrolytic etching of the fine wire after welding to the V support. The solutions used were designed to give what appeared to be a smooth finish to the point as viewed under the microscope. Tungsten and molybdenum were etched in caustic soda followed by ammonia.* Nickel was etched in a solution of potassium perchlorate in dilute hydrochloric acid. The etching process was continued, with repeated examination under the microscope, until the actual point could not be clearly resolved. The mean radius of point generally used was $2\cdot 5 \times 10^{-5}$ cm.

For tungsten, both specially prepared single crystal wires and polycrystalline wires were etched, but the patterns were such as to leave no doubt that in both cases a single crystal point was formed.† For molybdenum and nickel, polycrystalline wires were used.

(d) *Exhaust technique*

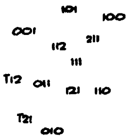
Early experiments showed that unless a very high order of vacuum was obtained, points were easily destroyed, and the patterns that were obtained showed changes with time which were obviously caused by the presence of gas. We shall refer to these effects in detail later in the paper. To overcome these difficulties the bulb and re-entrant foot tube were made of a special hard glass which could be baked at 700° C.

Pumping was carried out on a three-stage mercury diffusion pump fitted with a McLeod gauge reading to 10^{-6} mm. Hg and liquid air trap. The tube was baked for several hours without liquid air applied to the trap until the

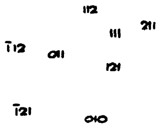
---

* This technique is derived from early industrial practice in reducing the diameter of drawn tungsten filaments.

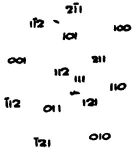
† There has been one exception, a thoriated tungsten wire which is discussed in a later paper.



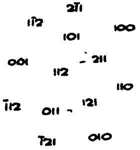
A FRESHLY ETCHED POINT



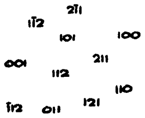
B PATTERN SOMETIMES OBTAINED FROM FRESHLY ETCHED POINT



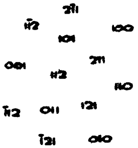
C POINT SMOOTHED BY HIGH TEMPERATURE FLASH



D POINT HEATED AT 1170°K — 1230°K WITH FIELD APPLIED



E POINT HEATED TO 1470°K WITH FIELD APPLIED



F POINT HEATED TO 1570°K WITH FIELD APPLIED

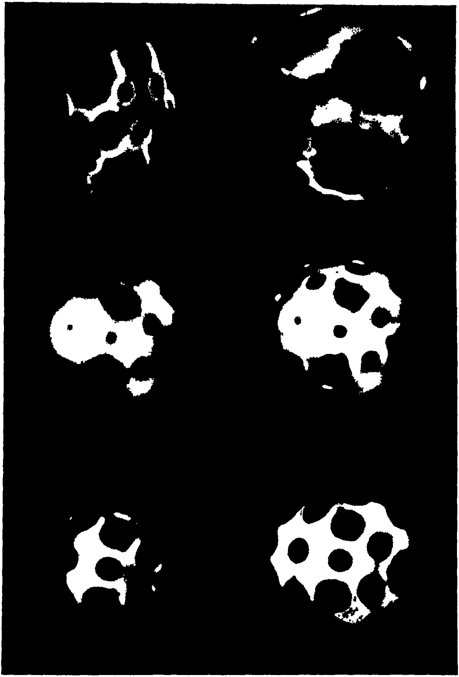


FIGURE 3 Field emission patterns from tungsten.

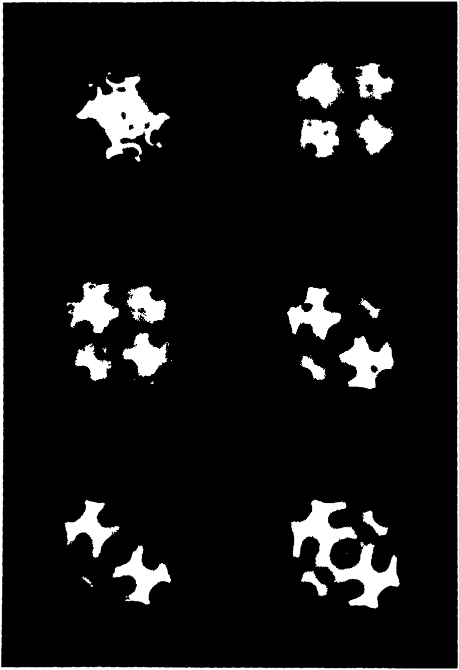
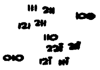
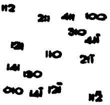


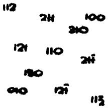
FIGURE 5*a*-*f*. Field emission from molybdenum showing build-up after heating at various temperatures with applied field



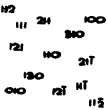
A UNFLASHED POINT



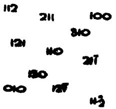
B FLASHED POINT



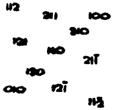
C POINT HEATED TO 770°K



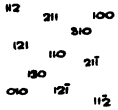
D POINT HEATED TO 920°K



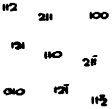
E POINT HEATED TO 1020°K



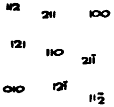
F POINT HEATED TO 1070°K



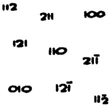
G POINT HEATED TO 1170°K



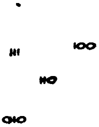
H POINT HEATED TO 1270°K



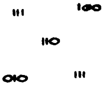
I POINT HEATED TO 1320°K



J POINT HEATED TO 1520°K



K POINT HEATED TO 1770°K APPEARANCE WHEN FIELD FIRST APPLIED



L POINT AT 1770°K FOR FOR SOME MINUTES WITH FIELD APPLIED

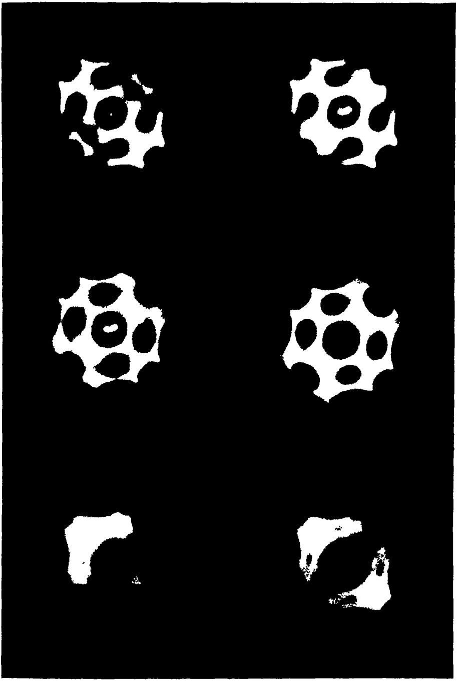


FIGURE 5*g–l*

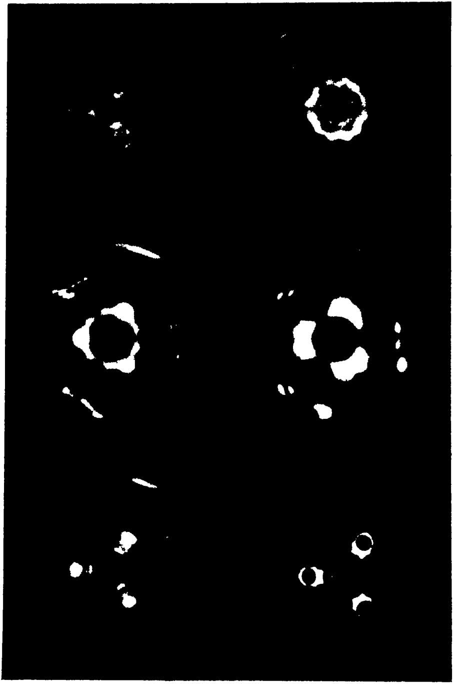
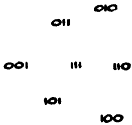
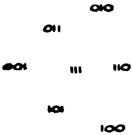


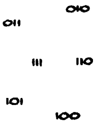
Figure 6*a–f*. Field emission patterns from nickel.



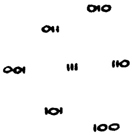
A FLASHED TO GET SMOOTH SURFACE H T 6000V



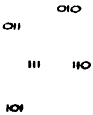
B HEATED TO 870°K H T 6000V



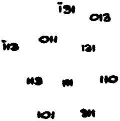
C HEATED TO 970°K FOR SEVERAL SECONDS H T 6000V



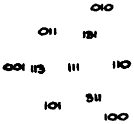
D HEATING CONTINUED AT 970°K FOR 5 MINUTES H T 5000V



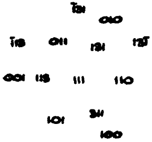
E REPEAT RUN, POINT SMOOTH H T 6000V



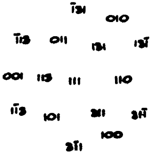
F HEATED TO 770°K FOR 1 MINUTE H T 6000 V



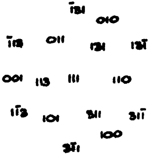
G 770°K – 5 MINUTES H T 5000V



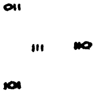
H HEATED TO 820°K H T 5000V



I HEATED TO 870°K FOR 1 MINUTE H T 5000V



J HEATED TO 920°K FOR 1 MINUTE H T 5000V



K HEATED TO 1100°K FOR 3 MINUTES H T 4000V

L HEATING MAINTAINED AT 1120°K AND H T INCREASED TO 5000V

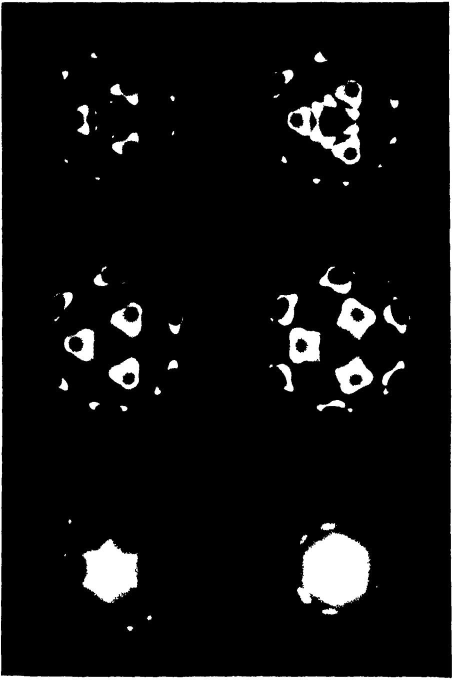


FIGURE 6*g*–*l*

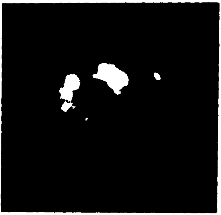


FIGURE 8. Build up on nickel just prior to flash-over

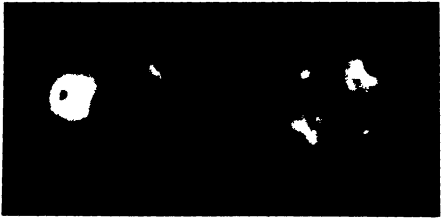
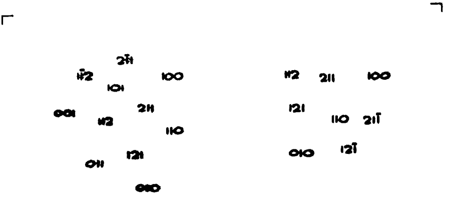


FIGURE 9. Patterns from tungsten and molybdenum after standing in relatively poor vacuum



A TUNGSTEN

B MOLYBDENUM

McLeod gauge was registering its minimum pressure. The getter was thoroughly outgassed by running it just below the reaction temperature, and the tungsten point was heated to 2800° K to remove any occluded gas. The tube was then rebaked, and finally liquid air was applied to the trap, the getter dispersed and the tube sealed off.

3. THE TUBE AS A SIMPLE MICROSCOPE

The field at the surface of the metallic point is given approximately by $V/2{\cdot}5\times10^{-5}$ V/cm., where V is the potential difference measured in volts between anode and point. Thus when V is 5000 V, the surface field is 2×10^{8} V/cm., which is sufficiently high for an appreciable electron current to be drawn out of any cold metal surface of work function 5 eV or less. If we assume that the electrons leave the point normal to its surface and continue to travel from point to screen in straight lines, the device acts as a simple projection microscope whose linear magnification is given by (radius of bulb)/(radius of point), which is in our case, approximately 200,000. A local disturbance of surface field such as might be produced by a single atom would at this magnification result in a patch of light or dark roughly of the order of $\frac{1}{10}$ mm. in diameter, assuming no diffraction. If the assumption is retained that the electrons are moving radially, a rough estimate of the size of spot due to diffraction effects can be made by assuming a parallel electron beam is emerging from an aperture of atomic dimensions. Then, using the treatment of the well-known optical case, the diameter of the spot due to diffraction on the screen is given by $\lambda R/d$, where R is the radius of the bulb, d is the diameter of the aperture and λ is the electron wave-length. The total spot size, including the magnification factor, is given by $D = R(\lambda/d + d/r)$. When V is 5000 V, $\lambda = (150/5000)^{\frac{1}{2}}\times10^{-8}$ cm. $= 1{\cdot}7\times10^{-9}$ cm. For minimum size of spot, $d^2 = \lambda r$, $d = 2{\cdot}1\times10^{-7}$ cm., or d should be about 5 atomic diameters. D is then 0·8 mm. The distance apart on the point for the resolution of two separate spots is then $0{\cdot}8r/R$ $= 4\times10^{-6}$ mm., or about 10 atomic diameters. We have plotted the spot size D as a function of the number of atoms across the diameter of a circular group causing the observed spot assuming the distance between adjacent atoms is 4×10^{-8} cm. This curve is shown in figure 2.

We have given this approximate theory because later we shall refer to the appearance of spots on the pattern whose diameters are of the order of 1–2 mm., which we shall treat as being images of small groups of atoms of this simple projection type formed by the tube acting as an electron microscope.

4. THE CALCULATED SIZE OF POINT

With tungsten points and tube made as described, electron emission patterns of high intensity could be obtained on the screen with anode voltages of 3000–10,000 V. Total emission currents were of the order of 1×10^{-5} amp. The actual voltage required for a pattern of a given intensity depended on the size of the point.

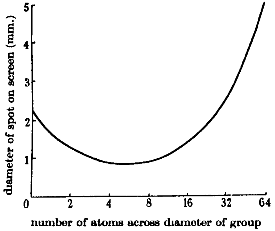


FIGURE 2. Size of observed spot against diameter of actual group of atoms.

The Fowler-Nordheim formula for field emission states that the electron current is given by

$$\frac{I}{A} = 6{\cdot}2 \times 10^{-6} \frac{\mu^{\frac{1}{2}}}{(\chi+\mu)\,\chi^{\frac{1}{2}}} F^2 e^{-6\cdot8\times10^7\chi^{\frac{3}{2}}/F}, \tag{1}$$

where I is the current in amperes, A is the emitting area in cm.^{2}, μ is the Fermi critical kinetic energy of the electrons in the metal, χ the work function in electron volts, and F is the field at the point surface in volt/cm. F can be written as $\beta V/r$, where β is a constant dependent on the actual shape of the point. V is the applied voltage and r the point radius. The formula fortunately is highly insensitive to μ. Writing $A = kr^2$, $\chi = 4{\cdot}5$ eV and $\mu = 8$ eV, we have $\log_{10} I/V^2 = \log_{10}[6{\cdot}6 \times 10^{-7}(k\beta^2)] - 2{\cdot}84 \times 10^8\, r/(\beta V)$. A plot of $\log_{10} I/V^2$ against $1/V$ should give a straight line whose slope gives r/β and whose intercept on the y-axis gives $k\beta^2$. Elimination of β then gives $kr^2 = A$ assuming that an accurate value of χ has been used.

We have made a number of such plots, and obtained excellent straight lines. Two such plots for molybdenum are discussed in detail later. The general finding is that r/β lies between 10×10^{-5} cm. and 5×10^{-5} cm.

From the study of equipotentials produced by a model of the point and filament system on a rubber membrane, we found $\beta = \frac{1}{2}$ approximately instead of the theoretical value unity for a smooth spherical point, because of the distortion introduced by the supporting V filament and wire. Then r lies between 5×10^{-5} and $2{\cdot}5\times10^{-5}$ cm., which is of the same order as that indicated by microscopic examination.

5 THE ANGULAR DISTRIBUTION OF AUTELECTRONIC EMISSION. TUNGSTEN

The patterns obtained from a tungsten point are shown in figure 3 (plate 4) All photographs reproduced were taken with the point at room temperature unless otherwise stated.

3*a* shows a typical pattern obtained after the point has not been very strongly heated on the pump prior to taking autelectronic currents.

It will be seen, allowing for the effect of photographing a spherical object, that the pattern has a high degree of symmetry, and, as Muller has shown, the centres of the dark spaces can be identified with directions normal to the main planes for a cubic crystal. The plane indices are shown in the transparencies above each plate In a particular case, X-ray examination of the point showed that the electron pattern was correctly interpreted. The angles between normals of the main planes measured from the field emission patterns were, however, only 80 % of their correct values, probably due to field distortion, caused by the presence of the supporting filament.

Sometimes the pattern in 3*a* is replaced by that shown in 3*b*. We believe that the various main planes on this point are etched in a series of steps which give rise to the emission ridges seen in the photograph, because of the stronger fields at the edges of the steps.

If the point is now heated for some time at 2800° K, in the absence of the external field, a pattern of the type shown in 3*c* is obtained. The symmetry is the same as in 3*a*, but the emission is much more uniformly distributed, and the dark centres have decreased in size. The $\{111\}$ as in 3*a* does not appear as a dark centre, but, although the $\{111\}$ region is uniformly emitting, the maximum intensity of emission on the pattern is actually in the zone immediately surrounding the $\{100\}$ dark centre. The $\{111\}$ is next brightest, and the zone about the $\{110\}$ least bright.

If the point temperature is now slowly raised, with the external field applied, a series of remarkable changes is observed. At room temperature the pattern has the appearance shown in 3*c*; as the temperature rises the regions between the {211} and {110} centres break up into a series of small spots of about 1 mm. diameter which execute a random motion. This motion is first clearly apparent at 1170° K. At 1230° K the motion is very rapid, bright patches appear in the centres of the {211}–{110} regions, the dark centres of the {211} and {110} increase in size and the emission increases over the {111} zone (3*d*). As the temperature is raised further, these bright patches become more intense and the motion is so fast that it can no longer be clearly seen (1470° K). At this stage, however, motion is seen on the outer parts of the {100} surface, and this motion spreads inwards towards the {100} centre until the whole of the {100} appears as a twinkling mass of spots of about 1 mm. diameter. At this stage too, the bright patches on the centres of lines joining the {211}–{110} begin to spread round the {110} and towards the {100}. Also the {100} dark centre begins to increase in size (3*e*). At 1570° K, the spread around the {110} and {100} is complete, leaving the planes as large dark centres outlined by intense emitting regions (3*f*). At higher temperatures, thermionic emission from the supporting system becomes sufficiently great for the applied field to be maintained no longer.

At any stage in the process, the pattern can be 'frozen' by lowering the point temperature to room temperature with the field still applied. In this way we have been able to obtain figures 3*d* and 3*e* showing the intermediate stages, and 3*f* showing the final stage of the changes described.

If, at 1670° K, the field is removed for a few seconds, and the point temperature is then reduced to room temperature in the absence of the field, the pattern has returned to normal (3*c*), when viewed at room temperature. This process can be repeated any number of times.

6. DISCUSSION OF PATTERN CHANGES

Muller (1939), who observed the changes in tungsten to the state shown in 3*e*, has suggested that the migrating atoms reposition themselves to form new surfaces of different work functions.

We believe the observed phenomena can be interpreted on the following lines:

(i) The original point is a single crystal with the main planes sharply defined by etching. The suggested appearance of a section through the point is shown in figure 4*a*. Provided no intensive heat treatment of the point takes place, the surface remains unchanged, and the pattern has the appearance of 3*a*, because the field is most intense at the various edges of the surface. Such a surface will be termed an 'etched' surface, and its pattern is the 'outline' pattern of the surface.

Heating the point at a high temperature results in a surface mobility and the surface atoms rearrange themselves in a manner which is determined by the periodic forces due to the underlying networks of atoms which cause the surface tension to be a function of crystallographic direction, and the surface atoms form a new surface of minimum free surface energy which is very much smoother, but shows departures from the perfect smoothness due to surface forces in a liquid. We suggest that a section through the point may appear as in 4*b*. The pattern is then that shown in 3*c*. It should be noted that there is no sharp edge to the pattern, and the very wide solid angle of 120–130° indicates that the field is nearly uniform over the greater part of the surface. This suggests that the shape of the point is as indicated in 4*c*. Where the field may be weakest, over the flat portions shown in 4*b*, dark centres occur in the pattern. (This is discussed further in a later section.) The variations in emission intensity over the bright parts of the pattern are probably due to differences in work function, since the point is in the smoothest state obtainable, after heat treatment in the absence of field. We would place the work functions of the different regions, in order of increasing work function, as $\{100\}$, $\{111\}$, $\{110\}$. This order is not quite in agreement with the measurements of Nichols (1940) obtained by thermionic methods from a tungsten single crystal. He obtained the order $\{111\}$, $\{100\}$, $\{110\}$.

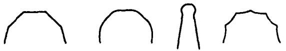


FIGURE 4. Diagram of sections through the point. *a*, point as etched; *b*, point after flashing; *c*, section to give wide angle emission; *d*, built-up point.

Recently, Martin (1939) has examined the low magnification image produced by thermionic emission from a single crystal and has given the same order as ours for the work functions. It is of interest to note that his thermionic emission pattern bears a considerable resemblance to the field emission patterns. Although he notes the thermionic emission decreases over the region about the $\{100\}$ as one goes towards its centre, he makes no mention of any non-emitting central spot.

When the point temperature is raised the surface atoms become mobile, and mobility commences first over the $\{211\}$–$\{110\}$ region, then over the

{111} and finally over the {110} region. In the presence of the external field, which is of the same order as the field due to the underlying surface atoms themselves, rearrangement of the surface atoms occurs, and the atoms pile themselves up in new equilibrium positions at certain points, so that a series of projections occur on the surface. The external field is now much higher at these projections than in the hollows, and these projections have much greater field emissions.

A section through such a point may appear as in 4*d*. The pattern passes through the stages 3*d*, 3*e* and 3*f*. On removing the field, at high temperatures, the surface atoms, which are mobile, are pulled back into the normal equilibrium position by the underlying surface fields to produce again the surface of minimum free energy.

If, however, the field is maintained while the temperature is reduced, the lack of surface mobility at lower temperatures (below 1170° K for tungsten) results in the surface remaining frozen into the 'built-up' form.

7. ANGULAR DISTRIBUTION OF AUTELECTRONIC EMISSION FROM MOLYBDENUM

In figure 5 (plates 5, 6) are shown the patterns from molybdenum. It will be noted that the etched surface produces a pattern (5*a*) similar to that from tungsten. The smooth surface (5*b*) shows very clearly the {310} planes, and is characterized by the four dark bands radiating from the {110}, and separating the {111} region from the {100} region.

On raising the temperature, motion about the {211}–{110} region is first observed at 770° K. At 870° K bright patches appear on lines joining the {211}–{110} centres, and the emission spreads over the {111} (5*c*).

At 920° K movement appears on the {111} zone, and the atoms appear to recede from the {111} centres, leaving them dark, and a spread of emission begins to occur around the {110} dark centre (5*d*).

With further heating, the surface atoms continue to move vigorously over the surface. At 1020° K the {111} dark centres have disappeared, and the {211}'s have become larger and elongated. The {310}'s now appear as large dark centres (5*e*). Figure 5*e* was taken with the point at 920° K, and the bright centre of the {110} is the filament loop which also appears in the subsequent photographs, all taken between 920 and 1070° K. A doubling effect is also noted at this stage. At 1070° K emission begins to spread around the {100} zone, which is now in motion, and the doubling effect is disappearing (5*f*). 5*g* shows the emission pattern at 1170° K. The {211} and {310} are now joined by dark bands. At 1270° K (5*h*) the build-up about the {100} continues, and the {310}'s have almost disappeared as dark centres due to heavy build-up. At 1320° K (*5i*) the build-up of emission about the {100} continues, and narrow dark bands join the {211}–{100} centres. At 1520° K the {100} is almost completely outlined by a ring of intense emission (*5j*). Note that the {111} area is now narrower and the {110} and {211} centres are much larger than in *5b*. At 1770° K a sudden instability occurs with high fields and the surface flows so as to concentrate the built up area around the {111} plane (*5k*). On standing at this temperature with field applied the pattern becomes symmetrical, as shown in *5l*.

Above 1770° K the thermionic emission from the supporting loop renders further observation impossible. The impression obtained throughout is that of flowing surfaces piling up at particular regions. If at 1770° K the field is removed for several seconds and the pattern then viewed at room temperature, the appearance is that of the smooth point of *5b*. The time taken for the surface to return to normality (*5b*) in the absence of the field is dependent on the actual point temperature. In table 1 we record the times taken to return to normal at different temperatures, after the surface had been completely built up.

TABLE 1. TIME FOR MOLYBDENUM SURFACE TO RETURN TO NORMAL

Temp. °K	Time sec.	Appearance (figure no.)
1570	15	*5d*
	45	*5c*
	90	*5b*
1470	15	*5e*
	120	*5c*
	210	*5b*
1370	15	*5g*
	480	*5d*
	1080	*5b*
1270	1080	*5c*
	2280	*5c*
1170	4140	*5d*
	7740	*5d*
1070	After 7740 sec.	*5j*
290	After several days	*5j*

It is evident that below 1170° K, the surface mobility is such that the surface remains indefinitely in the new state in the absence of field.

8. Distribution of autelectronic emission from nickel

The patterns from nickel are shown in figure 6. With nickel we cannot say with certainty what represents the truly smooth state, because even at room temperature, some slight 'piling-up' occurs, and here we observed an extremely well marked dependence of the pattern changes on the value of the applied field. The pattern can be modified or returned to normality by increasing or reducing the applied field.

Figure 6*a* (plate 7) represents our nearest approach to smoothness; it is symmetrical about the {111} which appears as a large dark centre. Groups of atoms can be seen around the {111}. Nickel has a face-centred structure, and may etch differently from a body-centred crystal. Figure 6*b* (plate 7) is after heating at 870° K with applied field. (The values of applied voltage are given in the reproductions.) Figure 6*c* (plate 7) is after heating at 970° K for a few seconds and 6*d* after further heating at 970° K. Note in 6*d* the appearance of nickel crystallites around the {100} caused by the surface piling up.

If the field is maintained for some time or increased slightly the build-up at these crystallites is such that the field becomes very high and the crystallite can be pulled off. Figures 6*e*–6*l* (plates 7, 8) show a repeat run on the same point after it has been flashed at 1270° K with no applied field, to restore it to normal The whole surface seems to build-up round the {111} and then finally piles up on to it. The current increases, and a further slight increase in field will pull off the point. Removing the field results in return to the normal state.

The nickel surface is visibly mobile at all temperatures above 370° K and an increase of applied field at any temperature can build up the pattern to the stage where the point can be pulled off.

9. The changes in pattern with temperature and applied field and the geometrical shape of the point

The results obtained with molybdenum and nickel leave no doubt that as the temperature of the point is raised the surface atoms become mobile, and in the presence of the applied field, the surface can be modified.

We have suggested that the shape of the point undergoes a change from a nearly smooth spherical state to a state which is roughly spherical, with hollows and peaks which vary according to the temperature and the strength applied field. The changing patterns obtained are then due to the variations in surface field caused by the surface alterations. A study of the variation of total emission with applied field for the different states ranging from the almost smooth state to completely built up state, gives additional support to our views.

In the cases of the smooth state and the completely built up state for molybdenum which gives the most stable sets of patterns, the plots of I/V^2 against $\log_{10}(I/V)$ (figure 7) show a marked decrease in slope (which does not involve the total emitting area) as between the smooth and built up

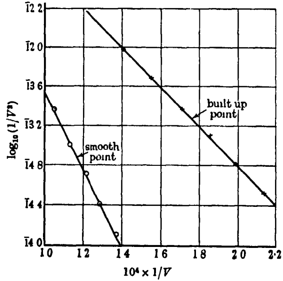


FIGURE 7. Field plots from smooth and built-up molybdenum points.

states. The slopes are $3\cdot94 \times 10^4$ and $2\cdot06 \times 10^4$ respectively. If χ_1, χ_2 are the respective mean work functions, and β_1 and β_2 are the values of β in each case, then

$$(\chi_1/\chi_2)^{\frac{3}{2}}\,(\beta_2/\beta_1) = 3\cdot94/2\cdot06.$$

If β is constant, then $\chi_1/\chi_2 = 1\cdot54$. If χ_1 assumed to be about 4·5 eV, then χ_2 is 2·92 eV. Now the measured maximum and minimum values of work function for tungsten (Nichols 1940) are 4·66 {211} and 4·35 {611} and similar values can be assumed for molybdenum, so that a simple change in effective work function does not appear to account for the observed differences. On the other hand, if χ is assumed constant, β_1/β_2 is almost 1/2, and this means a 2 : 1 change in field at the points of intense emission, which are the points mainly responsible for the total emission values of the plot.

A change in work function from 4·66 to 4·35 would, from the plots, give only a 2 : 1 change in emission current for a constant field, whereas a 2 : 1 change in β would increase the emission by nearly 100 : 1, and this change accords more nearly with the observed changes in emission intensity.

If one assumes a surface similar to that shown in figure 4*d*, the condition necessary for the field to change by 2 : 1 is that the height of the peak shall be about 25 atomic diameters above the hollow, where the surface is represented by a spherical harmonic of the 12th order *

The calculated area for the smooth point is $1{\cdot}7 \times 10^{-9}$ cm.^{2} and for the built up point is $8{\cdot}6 \times 10^{-11}$ cm.^{2}, a decrease of 15 : 1. This agrees with the observed fact that the bulk of the emission comes from only a few regions on the built up point.

10. OTHER ASSOCIATED PHENOMENA

(*a*) *The flash-arc phenomenon*

It is well known that in high vacuum devices using high voltages between electrodes, flash-overs often occur at welds and other rough surfaces, and are influenced by the presence of gas (Gossling 1932). The explanation is, we believe, that the welds are associated with fine points, and consequently high surface fields. When the metal gets hot, because of the actual passage of current or due to other causes, build up occurs, and high currents are drawn from localized sections of the point surface. The heating effect may be sufficient to cause evaporation and ionization and the discharge occurs. It is of interest to note that in practice the flash-over is more prone to occur with nickel than with molybdenum, and tungsten is safer than molybdenum. These results accord with our observations on the temperature at which mobility and build up occurs. The appearance of nickel crystallites just prior to blowing-up is shown in figure 8 (plate 9). Further, if any gas of an electropositive nature is present, the ions will be returned to the built-up sections on the point surface, and will still further enhance the local emission, and increase the ease with which a flash-over occurs. We have repeatedly observed this latter effect in some of our early tubes where the vacuum was not of an extremely high order.

It is of interest to note that with points of size 3×10^{-5} cm. radius, and 5000 V applied between anode and point, the electrostatic force on the surface of the points assuming $\beta = \frac{1}{2}$ is about 20 tons/in.^{2}. This force is approaching the cold tensile strength of the metals used. Typical figures are tungsten 300 tons/in.^{2}, nickel 75 tons/in.^{2} and iron 50 tons/in.^{2}. At high temperatures the tensile strengths are less. In the case of nickel the tensile strength at 1070° K is about one-quarter of the cold value. It is thus clear that there is a very good chance of pulling the point off, and this is probably the reason why it was not possible to go much above 1070° K with nickel, and why iron points present so many difficulties

---

* We are indebted to Dr R. W. Sloane for these calculations.

(*b*) *Effects of gas on the pattern*

In our earlier experiments the glass spherical bulb used was made of a lead glass of low softening point (450° C), so that high temperature bakes could not be employed. We had found in such tubes, that on running with the point at room temperature, the pattern from tungsten or molybdenum slowly undergoes modification

The altered patterns are shown in figure 9*a* for tungsten and 9*b* for molybdenum (plate 9). Note the difference between 9*a* and 3*c*, and 9*b* and 5*b*. On standing, dark bands have appeared joining the {211}–{110} centres, and at the same time the {211} centres have increased in diameter. On raising the point temperature to 620–670° K, the pattern immediately returns to normal. If the tube is immersed in liquid air, the effect can be prevented. We infer that although the pressure in the tube is of the order of 10^{-7} mm. Hg, sufficient gas is still present to settle down on parts of the surface. On immersing in liquid air, the pressure is considerably reduced. As Langmuir has indicated, immersion in liquid air with a barium getter present results in pressures of the order of 10^{-14} mm. Hg. The observations suggest that the gas film is of a type which is readily removed from the surface at 620–670° K, and that although the whole point surface is being bombarded with gas particles, there are preferred zones on which the electronegative gas atoms stick tightly.

We shall have more to say about the preference of certain types of contaminating atoms for particular regions on the surface in the second paper on contaminated surfaces. Here, however, we would point out that the increase in size of the {211} dark centres under conditions where no mobility of surface atoms of the metal itself is suspected, lends some support to the suggestion made later that the dark centres in the pattern from the 'clean' smooth points may be due to the presence of negative ions which cannot be removed below the melting point, and which result in very high work functions over the small flats present on the point surface in its smoothest form.

A similar effect is observed, when contaminations are liberated by intense bombardment of the getter or the screen itself.

The pattern can become covered with a mass of disappearing and reappearing bright spots due to electropositive contaminations. Here, again, these can be removed or prevented by holding the point at 620–670° K.

11. THE PATTERNS FROM THE SMOOTH POINTS

The patterns for smooth tungsten and molybdenum show first a variation in emission density around the main planes which is never more than 10 : 1. The dark centres are indicative of at least a 100 : 1 change relative to the adjacent bright areas. If work function changes are assumed to cause the dark patches in the pattern from smooth surfaces, then the variations must be of the order of 2 : 1, which is very much greater than any differences of work function for different faces measured experimentally. On the other hand, a surface such as we have suggested for the smooth point (4*b*), does not appear to us to permit of a field difference of 2 : 1 from the edge of the small flat to its centre. If we combine the known work function variation with a field increase, we still require almost a 2 : 1 increase in field.

There exists also the possibility that an electron optical effect, due to the shapes of the equipotentials close to the point surface, may cause electrons leaving the flat part to be deflected in such a way as to reinforce the emission coming from the edges and so cause a dark centre to appear on the screen To test this possibility, a steel sphere of 5 mm. diameter with a flat ground on its surface, was mounted in the centre of a large spherical bulb. The surface of the sphere was coated with barium oxide, and its thermionic emission pattern was thrown on to the fluorescent screen on the inside of the bulb. The pattern showed a bright circular patch, opposite the centre of the flat, surrounded by a faint dark ring, and the rest of the sphere gave rise to a uniform emission, somewhat lower in intensity than that due to the centre of the flat portion. Thus it would appear that any electron optical effects should show some emission in the middle of the dark centres. Even if the surface irregularities are slight hollows rather than flats, electron optical effects would only modify the pattern to a very small extent.

One explanation would be to assume that in a surface such as that shown in 5*b*, the metal cannot be completely freed from electronegative gas contamination, and that on the small flats, where the van der Waals forces will be greatest, electronegative gas is firmly held as atoms or negative ions to the tungsten. There is some slight support for this view, as will be seen from the description of the pattern changes in the presence of gas which were discussed in § 10. With tungsten, such a contamination would be oxygen, and then the work function would increase from 4·52 to 9·2 eV, a change of just the order required to explain the differences in emission intensity observed.

Although we have suggested this last possibility, we have no reason to believe that oxygen cannot be removed from tungsten at temperatures above 2000° K. On the other hand, it may be that the failure to obtain a perfectly clean surface is one of the reasons why, in the thermionic emission formulae, the constant A, for both tungsten and molybdenum, always is less than the theoretical 120 amp /cm.^{2} degree^2.

In the case of nickel, we would then have to suppose that the electronegative gas can attach itself firmly to the $\{111\}$, which appears as a small flat in the face-centred structure, but which does not appear in the body-centred structure.

There is an aspect of the problem which the evidence does not permit us to ignore. There exists the possibility that the electron pattern is a true representation of the electron kinetic energy distribution inside the clean smooth metal, and that according to the structure, the electron energies are a minimum along the normals to certain main planes, which are different for the body-centred and face-centred types of lattice. This means a variation in work function over the crystal surfaces, the work function rising steeply along certain of the main crystallographic directions.

We have made an attempt to secure further data by using iron as the point metal. Iron changes from the body-centred to the face-centred form at a relatively low temperature, but so far we have not succeeded in obtaining stable patterns.

An alternative explanation would be that the electron waves for electrons at the top of the Fermi surface are reflected back for perpendicular incidence on certain planes. We make this point because of the connexion that we find exists between the field emission patterns and X-ray reflexion patterns. In the case of a body-centred cubic crystal (W, Mo) the condition for the $\{hkl\}$ plane to give an X-ray reflexion is that $(h+k+l)$ shall be even. The planes with the largest lattice spacing to fulfil this condition are $\{110\}$, $\{200\}$, $\{211\}$, and $\{310\}$, and in the field emission pattern it is these that appear as dark centres. We cannot distinguish between the various orders.

Again, for a face-centred cubic crystal (Ni), the condition for X-ray reflexion is that h, k, l must be all odd or even. Thus the $\{111\}$, $\{200\}$, $\{220\}$ and $\{311\}$ planes appear. The smoothest nickel pattern (6*a*) shows the $\{111\}$ plane as a dark centre. We have attempted to connect the electron wavelength associated with μ, the kinetic energy of the electrons at the top of the Fermi distribution which are those that dominate the field emission, with the condition for Bragg reflexion, given by

$$\frac{a_0}{\sqrt{(h^2+k^2+l^2)}} = n\frac{\lambda}{2}, \tag{2}$$

where a_0 is the lattice constant, and λ is the incident wave-length, but the values of μ obtained lie outside the possible values for tungsten, molybdenum and nickel except in the case of the {110} plane for tungsten and molybdenum, and the {111} or {200} for nickel.*

There is one aspect of the problem to which we would like to draw attention. The structure-amplitude expression which determines the planes that take part in the X-ray reflexion also determines the form of the Brillouin zones in momentum space. This factor is $S_{hkl} = 1+\cos\pi(h+k+l)$ for a body-centred cubic lattice, and

$$S_{hkl} = 1+\cos\pi(k+l)+\cos\pi(h+k)+\cos\pi(h+l)$$

---

* Since writing the above, our attention has been drawn to a note by Muller (1939), in which he has attempted to correlate the dark centres of the tungsten pattern with the Bragg reflexion condition. Muller uses this in the form

$$\frac{ma_0}{\sqrt{(h^2+k^2+l^2)}} = n\frac{\lambda}{2}, \tag{3}$$

where ($m, n = 1, 2, 3, 4, \ldots$). By putting $\mu = 7·5$ eV, corresponding to 1·5 electrons per atom for tungsten, he obtains an electron wave-length of 4·5 A ($\lambda = (150/\mu)^{\frac{1}{2}}$ A), which satisfies the equation (3) for the {110} plane, $m = 1$, $n = 1$ Values of m and n can be found to satisfy approximately the equation for other planes. Thus for the {211}, $m = 7$, $n = 4$ and for the {100} $m = 5$, $n = 7$ It appears to us, however, that the introduction of m is not justified except in special cases, and then when only one or two plane networks are involved In the case of a pure metal, the introduction of m would mean the appearance of fractional orders in both X-ray and electron diffraction spectra, and one would not expect the sharp edges to the dark centres obtained in the field emission patterns. The values of μ obtained from equation (2), for different planes are set out in the following table.

	Plane	μ eV ($n=1$)
Tungsten	{110}	7·48
	{200}	13·96
	{211}	22·44
	{310}	37·40
Nickel	{111}	9·06
	{200}	12·08
	{220}	24·16
	{311}	33·22

Values of μ for tungsten are 5·8 and 9·2 eV corresponding to one and two electrons per atom respectively, and for nickel $\mu = 11·7$ eV, for two electrons per atom. If one assumes μ may have a single value which is greater than 5 eV and less than 15 eV, it will be seen that only one plane in each case will satisfy the equation (2).

for a face-centred cubic lattice. Thus the planes which appear as dark centres might be the bounding planes to the Brillouin zones. The electrons at the top of the Fermi distribution can be represented by points lying on a surface of constant total energy. It would only be necessary for this surface to touch or nearly touch the bounding planes of the zones in order that the value of the kinetic energy of an electron may be such that it will make a Bragg reflexion at a plane with the same indices in the crystal lattice.

If the present phenomena are explicable in terms of Brillouin zones, they imply equality between the energies at several bounding planes of a single zone. The information at present available leaves the correctness of this explanation a matter for speculation, but if it be correct it would follow that for a monovalent face-centred metal like silver, the first Brillouin zone, bounded by {111} and {200} planes, would contain the Fermi surface entirely within it (the zone being half full). This would only allow at the most two dark centres to appear. Similarly a monovalent body-centred metal like sodium whose first zone is bounded by the {110} planes could only show one dark centre at the most corresponding to the {110} direction. These metals are so weak mechanically, that it may be very difficult, if not impossible, to obtain emission patterns (§ 9), but we are attempting the experiments in the hope of getting more evidence.

The authors wish to express their thanks to Professor R H. Fowler for many helpful suggestions and discussions

REFERENCES

Benjamin, M. and Jenkins, R. O. 1938 *Phil. Mag.* **26**, 1049.
— — 1939 *Nature, Lond* , **143**, 549.
Gosling, B S. 1932 *J. Instn. Elect. Engrs*, **71**, 461.
Martin, S. T. 1939 *Phys Rev* **56**, 947.
Muller, E. 1937 *Z. Phys.* **106**, 541.
— 1938 *Z. Phys.* **108**, 668
— 1939 *Naturwiss.* **49**, 821
Nichols, M. H. 1940 *Phys. Rev.* **57**, 297.

Two-phase equilibrium in binary and ternary systems
IV. The thermodynamic properties of propane

By J. H. Burgoyne

Department of Chemical Technology, Imperial College, London

(*Communicated by A. C. G. Egerton, Sec.R.S.—Received* 4 *June* 1940)

Existing physical and thermal data relative to propane have been summarized and correlated, and some new experimental determinations of pressure-volume-temperature relationships for the liquid at low temperatures have been carried out to make good deficiencies in the literature.

On the basis of the information thus obtained the entropy and enthalpy of propane have been calculated for conditions of temperature between −80 and 200° C, and at pressures of from 0·1 to 200 atm. The results are tabulated and also presented graphically on a temperature base.

Introduction

Recent developments in the utilization of propane have revealed the need for more complete information as to its thermodynamic properties and, in particular, the desirability of constructing an entropy diagram covering a wide field of temperature and pressure. On examining the literature it is found that sufficient thermal and physical data are available to achieve this object in a serviceable manner, although the accuracy is not everywhere all that is to be desired.

Certain entropy data, of restricted application, have already been published. Thus, Dana, Jenkins, Burdick and Timm (1926) have calculated the entropy of the saturated liquid and vapour from the normal boiling-point up to *c.* 50° C, whilst Sage, Schaafsma and Lacey (1934) have published a more comprehensive diagram extending from 20° C up to the critical temperature, with pressures ranging from 1 to about 200 atm. In both cases English units (B.Th. units, pounds per square inch, cubic feet, pounds and degrees Fahrenheit) were employed. A diagram of Justi (1936) appears to be identical with that of Sage, a centigrade scale being appended.

A temperature-enthalpy chart has been published by Ragatz (1934) covering all ordinary practical conditions of temperature and pressure, but lacking in detailed information.

The chief requirements, therefore, appear to be entropy and enthalpy diagrams which will correlate the work of Dana and Sage and will take into account additional data which have come to light since 1934. On these lines it is possible to cover the ranges 0·1 to 200 atm. and -80 to $+200°$ C satisfactorily. Data is lacking regarding the P-v-T relationships of the supercooled liquid at low temperatures, but this deficiency has been made good by new measurements. The units used in the present work are litres, kilogrammes, normal atmospheres, kilogram-calories and degrees Centigrade. A list of the symbols used is given at the end of the paper.

METHODS OF CALCULATION OF ENTROPY AND ENTHALPY

In order to construct a complete entropy (or enthalpy) diagram for any substance, it is sufficient to know (i) the functional relationship between pressure, volume and temperature in the states concerned, and (ii) the relationship between specific heat and temperature under a specified condition of pressure or volume.

From the first, by a graphical or algebraic operation, the connexion between the saturated vapour pressure and temperature can be determined, and, with the aid of the Clausius-Clapeyron equation, the latent heat-temperature relationship can be derived. By partial differentiation of the equation of state, all isothermal, isobaric and isometric derivatives necessary for thermodynamic calculations can be evaluated.

From the second, the value of the specific heat under any given condition of temperature and pressure (or volume) can be determined thermodynamically by utilizing the appropriate derivatives calculated from the equation of state.

In practice, such a procedure would prove tedious, and also inaccurate, since no equation of state has yet been discovered which is applicable both to the liquid and gaseous states with the necessary degree of accuracy. It is customary, therefore, to make direct use of experimental data additional to that already specified, and the precise method of calculation to be employed depends upon the nature of the information available. In the present instance the procedure was as follows.

From data relative to the saturated vapour pressure the relationship between the boiling-point and the pressure was first determined, thus establishing the location with respect to temperature of various isobars in the two-phase region of the thermodynamic diagrams.

The specific heat-temperature relationship for the saturated liquid was then plotted and used to integrate the expression

$$\left(\frac{\partial s}{\partial T}\right)_s = \frac{c_s}{T}. \tag{1}$$

Similarly, the enthalpy-temperature curve was determined by integrating

$$\left(\frac{\partial h}{\partial T}\right)_s = c_s + 0{\cdot}0242 v_s \left(\frac{\partial P}{\partial T}\right)_s, \tag{2}$$

in which the derivative $(\partial P/\partial T)_s$ was found from the vapour-pressure equation. The factor 0·0242 converts litre-atmospheres to kilogram-calories.

In the case of both entropy and enthalpy, the saturated liquid at atmospheric pressure was taken as the zero-point.

The latent heat-temperature equation was then derived from experimental data and utilized to determine the increase in entropy and enthalpy on isobaric vaporization, according to the equations

$$\Delta s = \frac{l}{T}, \tag{3}$$

$$\Delta h = l. \tag{4}$$

An empirical c_P-T equation for the vapour at atmospheric pressure was established and used to integrate the expressions

$$\left(\frac{\partial s}{\partial T}\right)_P = \frac{c_P}{T}, \tag{5}$$

$$\left(\frac{\partial h}{\partial T}\right)_P = c_P. \tag{6}$$

Having thus determined the properties of the vapour at atmospheric pressure and various temperatures, its properties at higher pressures were calculated by integrating along isothermal paths the expressions

$$\left(\frac{\partial s}{\partial P}\right)_T = -0{\cdot}0242 \left(\frac{\partial v}{\partial T}\right)_P, \tag{7}$$

$$\left(\frac{\partial h}{\partial P}\right)_T = 0{\cdot}0242 \left[v - T\left(\frac{\partial v}{\partial T}\right)_P\right], \tag{8}$$

the derivative $(\partial v/\partial T)_P$ and the specific volume being obtained from experimental P-v-T data.

Starting from the saturated liquid curves, the properties of the supercooled liquid at higher pressures were calculated by integrating the equations (7) and (8) along isothermals, as in the case of the superheated vapour.

Above the critical pressure, corresponding isobars, determined independently by isothermal integration from the saturated liquid line and the 1 atm. vapour line, were connected up.

RESULTS

The normal boiling-point, critical constants and saturated vapour pressure

Before attempting to derive a vapour pressure-temperature relationship to represent the available data, it is convenient to establish two fixed points, namely the normal boiling-point and the critical point. Several determinations of each have been reported, comprising in the case of the normal boiling-point the following values:

Authority	Boiling-point °C
Olszewski (1889)	−45·0
Meyer (1894)	−37·0
Hainlen (1894)	−37·0
Lebeau (1905)	−44·5
Burrell and Robertson (1915)	−44·1
Maass and Wright (1921)	−44·5
Dana *et al* (1926)	−42·12*
Francis and Robbins (1933)	−42·3
Sage *et al* (1934)	−42·6*
Hicks-Bruun and Bruun (1936)	−42·17 ± 0·05
Hartneck and Edse (1938)	−41·14† ± 0·01

* Calculated by Cox (1934).
† 42·14 according to our interpretation of the results.

The earlier determinations were probably vitiated by the impurity of the gas used, and are in serious disagreement. Recent values are more concordant, but Hartneck and Edse give a figure higher than that previously accepted, although their published results suggest good agreement with other workers. The value determined directly by Hicks-Bruun and Bruun has been adopted in the present work, being in close accord with the figure calculated by Cox (1934) from the vapour pressure data of Dana *et al.* and with our own interpretation of the most recent work of Hartneck and Edse.

The following values of the critical constants have been reported:

Authority	Critical temp. °C	Critical pressure atm.	Critical volume
Olszewski (1889)	97·0	44·0	—
Meyer (1894)	102·0	48·5	—
Hainlen (1894)	102·0	48·5	—
Lebeau (1905)	97·5	45·0	—
Maass and Wright (1921)	95·6	—	—
Sage *et al.* (1934)	100·1	43·77	4·31 lit./kg.
Beattie, Poffenberger and Hadlock (1935)	96·81 ± 0·01	42·01 ± 0·02	4·43 ± 1%

There are, again, serious discrepancies in these data, and it is unfortunate that the evidence of Sage and of Beattie should be in sharp disagreement, since we are almost exclusively dependent upon their work for the P-v-T relationships. The later results of Beattie have, however, been accepted by Sage (1940) in recent work on propane and are adopted in the present calculations.*

Several determinations of the saturated vapour pressure in various temperature ranges have been carried out, more or less information being available from most of the sources quoted in the table of normal boiling-points. On plotting all the results together on a temperature base, there is found to be some disagreement.

The equation

$$\log P = \frac{a(T+b)}{T} \tag{9}$$

is not strictly accurate over the whole temperature range, but if in the case of propane b is put equal to the normal boiling-point and the equation is solved for a at the critical point, the following forms are obtained which represent the mean of the various experimental values exceedingly well at all temperatures:

$$\log_{10} P = \frac{4\cdot319(T-231\cdot0)}{T}, \tag{10}$$

or

$$T = \frac{997\cdot7}{4\cdot319-\log_{10} P}. \tag{11}$$

In view of this result there is little justification, in the absence of precise experimental data, for using the more complex equation due to Nernst:

$$\log P = \frac{\lambda_0}{4\cdot571}T + 1\cdot75\log T - \frac{\epsilon}{4\cdot571}T + c \tag{12}$$

(cf. Burrell and Robertson (1915), Dana *et al.* (1926)). Corresponding vapour pressure-temperature values calculated from equation (11) are included in table 1.

Specific heat of the saturated liquid

Data relating to the specific heat of liquid propane are due to Dana *et al.* (1926) and to Sage and Lacey (1935), both of which sources refer to the saturated condition. They cover low- and high-temperature ranges respectively. On plotting the experimental values on a temperature base, it is

---

* *Footnote added in proof.* The work of Beattie has been substantiated by the results of Deschner and Brown (*Industr. Engng Chem.* 1940, **32**, 836) published since the present paper was communicated.

TABLE 1. ENTROPY AND ENTHALPY OF THE SATURATED LIQUID AND VAPOUR

Pressure atm.	Temp. °C	Saturated liquid		Saturated vapour	
		Enthalpy	Entropy	Enthalpy	Entropy
1	−42·17	0	0	101·60	0·4400
2	−24·8	9·41	0·0391	107·11	0·4325
3	−13·4	15·69	0·0646	110·29	0·4290
4	− 4·7	20·55	0·0829	112·45	0·4253
5	2·6	24·71	0·0980	114·31	0·4232
6	8·7	28·23	0·1105	115·73	0·4211
7	14·1	31·41	0·1215	116·91	0·4193
8	18·9	34·29	0·1313	118·09	0·4185
9	23·5	37·09	0·1406	118·89	0·4164
10	27·5	39·57	0·1487	119·77	0·4155
12	34·9	44·27	0·1638	120·97	0·4130
15	44·3	50·51	0·1833	122·61	0·4107
20	57·5	59·94	0·2116	124·14	0·4058
25	68·5	68·47	0·2361	124·43	0·4000
30	78·0	76·36	0·2580	123·20	0·3915
35	86·4	84·54	0·2801	120·45	0·3800
40	94·1	95·08	0·3081	114·13	0·3600
42·01	96·81	105·54	0·3352	105·54	0·3352

found that the Dana points are rather scattered, but that a curve can be drawn representing the complete data fairly well.

Using the smoothed c_s-T curve and commencing with $s = 0$ for the saturated liquid at atmospheric pressure, the entropy-temperature curve for the saturated liquid was constructed by integrating equation (1). Similarly, the enthalpy-temperature curve was determined from the same starting-point by integrating equation (2). Values of v_s are available from the data of Dana *et al.* (1926) and Sage *et al.* (1934) covering the entire range of temperature. Earlier data on the specific volume of the saturated liquid are found in the work of Meyer (1894), Lebeau (1905) and Maass and Wright (1921). The concordance of values from the various sources is sufficiently good for the present purpose, since the second term in equation (2) plays but a small part in determining $(\partial h/\partial T)_s$.

Calculated values of s and h for the saturated liquid are given in table 1.

Latent heat of vaporization

The latent heat of vaporization of propane has been determined by Dana *et al.* (1926) (−40 to +20° C) and by Sage, Evans and Lacey (1939) (40–75° C). The two sets of results when plotted against the temperature are found to be reasonably concordant and the general empirical relationship originally proposed by Osborne and van Dusen (1918) represents them very well. The exact form of the equation for propane is found to be

$$l = 12{\cdot}08(T_c - T)^{\frac{1}{2}} - 0{\cdot}2932(T_c - T), \tag{13}$$

from which the values of l used are calculated.

These figures can be checked using the Clausius-Clapeyron equation:

$$l = \frac{dP}{dT} T(v_2 - v_1), \tag{14}$$

where v_2 and v_1 refer respectively to the saturated vapour and the saturated liquid at $T°$ K, and are derived from the data of Dana and Sage. dP/dT is obtained from the vapour-pressure data. The values of l thus found are in fairly good agreement with the smoothed experimental values, the maximum deviation being about 2 %. As the experimental values of v_2 appear to be less reliable than those of l, the Clausius-Clapeyron values for the latent heat were not employed in calculation.

By means of the expressions (3) and (4) the entropy- and enthalpy-temperature curves of the saturated vapour were constructed from those of the saturated liquid.

The superheated vapour

The isobaric heat capacity of propane at atmospheric pressure has been determined by Sage, Webster and Lacey (1937) between the temperatures 20 and 170° C. Their results when plotted against temperature show a linear relationship and there seems to be every justification for extrapolating to the limits required in the present work. The corresponding equation is

$$c_{P=1\,\text{atm}} = 0{\cdot}2085 + 0{\cdot}000668T. \tag{15}$$

With the aid of this relationship the equations (5) and (6) can now be integrated, giving the s-t and h-t curves for the 1 atm. isobar. Taking the values $s = 0{\cdot}4400$ and $h = 101{\cdot}60$ when $T = 230{\cdot}9$ from table 1, the following practical equations are obtained:

$$s_{P=1\,\text{atm}} = 0{\cdot}4802 \log_{10} T + 0{\cdot}000668T - 0{\cdot}8491, \tag{16}$$

$$h_{P=1\,\text{atm.}} = 35{\cdot}65 + 0{\cdot}2085T + 0{\cdot}000334T^2. \tag{17}$$

Having thus determined the 1 atm. isobars, it is possible to calculate from them the properties of the superheated vapour at higher pressures by employing the isothermal relationships (7) and (8). Some of the necessary P-v-T data are provided by the work of Sage and of Beattie, but there is a gap below 40 atm. above the critical temperature. This has been filled by assuming ideal behaviour at the highest temperature considered (200° C) and low pressures. At low pressures the derivative $(\partial v/\partial T)_P$ then has the value $nR/P = 0{\cdot}0449/P$ kcal., and an isothermal interpolation can be made to meet the Beattie data at higher values of P. An isobaric interpolation connects up with Sage's data at the critical temperature, and it is thus possible to plot out the entire entropy-temperature field for the vapour from 1 to 200 atm. At high temperatures and low pressures, the positive and negative terms in equation (8) become equal, and $(\partial h/\partial P)_T = 0$.

In the case of the vapour below the critical temperature and pressure, the results thus obtained can be checked by an independent method employing the Joule-Thomson coefficients determined by Sage, Kennedy and Lacey (1936). In this case

$$\mu = \left(\frac{\partial T}{\partial P}\right)_h, \tag{18}$$

i.e. μ is the differential of the isenthalpic relation of T and P, and by integration the P-T intercepts of the constant enthalpy line can be determined. From these results corresponding values of the entropy can be derived by calculating isobarically from the saturated vapour line using the relationship

$$\left(\frac{\partial s}{\partial T}\right)_P = \frac{c_P}{T} = \left(\frac{\partial h}{\partial T}\right)_P \frac{1}{T}. \tag{19}$$

The values of enthalpy and entropy calculated from the Joule-Thomson coefficients are in good agreement with those calculated from the P-v-T data except near the saturated vapour line. Sage has suggested that this discrepancy is due to the influence of traces of oil in the apparatus in which the P-v-T relationships were determined.

It is unfortunate that there is disagreement between Sage and Beattie as to the location of the critical point. While the more recent work of the latter carries the greater weight, his P-v-T data are insufficient for the present purpose, particularly just below the critical temperature, and it is essential to make some attempt to fit in the Sage values. In the circumstances, this part of the thermodynamic diagram must be regarded as less certain than the rest, and further practical investigation in extension of Beattie's experiments is to be desired.*

The properties of the superheated vapour are summarized in tables 2–5.

---

* This investigation has since been carried out by Deschner and Brown (see previous footnote).

TABLE 2. ENTHALPY BELOW THE CRITICAL PRESSURE

Pressure	Temp. °C										
atm.	−42·17	−24·8	0	2 6	18 9	34·9	50	57·5	100	150	200
1	0*	—	117·51	—	—	—	137·87	—	159·90	183 69	209·26
2	—	9·41*	116·73	—	—	—	137·11	—	159·15	—	209·26
3	—	—	115·94	—	—	—	136 32	—	158 36	—	209·26
4	—	—	115·30	—	—	—	135 68	—	157·72	—	209 26
5	—	—	—	24 71*	—	—	135·06	—	157·10	183·10	209·26
6	—	—	—	—	—	—	134·39	—	156·43	—	209·26
7	—	—	—	—	—	—	133·72	—	155 76	—	209·26
8	—	—	—	—	34 29*	—	133 00	—	155·04	—	209·26
9	—	—	—	—	—	—	132·53	—	154·57	—	209·26
10	—	—	—	—	34·27	—	131·88	—	153·92	181·45	209·26
12	−0·03	9 36	—	24 83	—	44 27*	130 66	—	152·70	—	209·26
15	—	—	—	—	—	—	128 76	—	150·80	179·95	209·20
20	+0·16	9·51	—	24·89	34 25	44·16	—	58 94*	147·69	178 20	208·76
30	0·48	9·77	—	25 05	34·32	44·13	—	59·59	134·10	170·60	206·95
40	—	—	—	—	34 33	—	—	—	124·25	164·35	204·23

* Saturated liquid.

TABLE 3. ENTROPY BELOW THE CRITICAL PRESSURE

Pressure atm.	Temp. °C											
	−42·17	−24·8	0	2·6	18·9	34·9	50	57·5	78·0	100	150	200
1	0*	—	0·5033	—	—	—	0·5717	—	—	0·6350	0·6948	0·7519
2	—	0·0391*	0·4694	—	—	—	0·5378	—	—	0·6011	0·6620	0·7229
3	—	—	0·4502	—	—	—	0·5186	—	—	0·5819	0·6434	0·7049
4	—	—	0·4358	—	—	—	0·5042	—	—	0·5675	0·6298	0·6921
5	—	—	—	0·0980*	—	—	0·4925	—	—	0·5558	0·6190	0·6821
6	—	—	—	—	—	—	0·4828	—	—	0·5461	0·6100	0·6739
7	—	—	—	—	—	—	0·4744	—	—	0·5377	0·6025	0·6670
8	—	—	—	—	0·1313*	—	0·4670	—	—	0·5303	0·5960	0·6610
9	—	—	—	—	0·1311	—	0·4604	—	—	0·5237	0·5900	0·6557
10	−0·0008	0·0384	—	0·0974	0·1309	—	0·4543	—	—	0·5176	0·5848	0·6510
12	−0·0010	0·0382	—	0·0972	0·1306	0·1638*	0·4437	—	—	0·5070	0·5752	0·6428
15	−0·0012	0·0379	—	0·0968	0·1301	0·1632	0·4299	—	—	0·4932	0·5628	0·6322
20	−0·0015	0·0375	—	0·0962	0·1292	0·1622	—	0·2116*	—	0·4727	0·5452	0·6174
25	−0·0018	0·0371	—	0·0956	0·1286	0·1613	—	—	—	0·4490	0·5274	0·6051
30	−0·0021	0·0367	—	0·0950	0·1278	0·1604	—	0·2088	0·2580*	0·4288	0·5118	0·5944
35	−0·0024	0·0363	—	0·0944	0·1270	0·1595	—	0·2076	—	0·4148	0·4997	0·5848
40	−0·0027	0·0359	—	0·0939	0·1262	0·1583	—	0·2064	0·2546	0·3944	0·4854	0·5762

* Saturated liquid.

TABLE 4. ENTHALPY ABOVE THE CRITICAL PRESSURE

Pressure	Temp. °C									
atm.	−42·17	−24·8	+2 6	18·9	34·9	57·5	78·0	100	150	200
50	0 96	10 16	25·31	34·39	43·96	59·12	75·15	101·90	157·25	201·80
60	—	—	—	34·45	—	—	—	96·40	148·80	197·64
80	—	—	—	34·63	—	—	—	93·65	140·40	190·72
100	—	—	—	34·85	44·33	58 76	73 09	92 00	135 66	183 65
120	—	—	—	35·13	—	—	—	—	132·06	177·93
140	—	—	—	35·40	—	—	—	—	129·92	173·84
160	—	—	—	35·72	—	—	—	—	128·15	170·66
180	—	—	—	36·04	—	—	—	—	126·94	168·37
200	—	—	—	36·36	45·66	59·72	72 75	88·70	126 06	166·73

TABLE 5. ENTROPY ABOVE THE CRITICAL PRESSURE

Pressure	Temp. °C									
atm.	−42·17	−24·8	+2 6	18·9	34·9	57·5	78·0	100	150	200
50	−0·0034	0·0350	0 0928	0 1248	0·1567	0·2042	0·2512	0·3230	0·4615	0·5650
60	−0·0040	0·0342	0·0918	0·1234	0·1552	0·2022	0·2476	0·3065	0·4350	0·5476
80	—	—	—	0·1208	0·1525	0·1985	0·2418	0·2920	0·4082	0·5234
100	—	—	—	0·1184	0·1500	0·1952	0 2372	0·2850	0·3932	0·5016
120	—	—	—	0·1162	0·1476	0·1923	0·2333	0·2780	0·3808	0·4842
140	—	—	—	0·1140	0·1453	0·1897	0·2298	0·2728	0·3722	0·4711
160	—	—	—	0·1120	0·1431	0·1873	0·2267	0·2685	0·3647	0·4605
180	—	—	—	0·1100	0·1410	0·1851	0·2239	0·2652	0·3587	0·4523
200	—	—	—	0·1080	0·1390	0·1831	0·2213	0·2620	0·3536	0·4455

The supercooled liquid

The properties of the supercooled liquid can be determined from those of the saturated liquid, given a complete knowledge of the *P*-*v*-*T* relations of the liquid under relevant conditions. Unfortunately, the published data are not complete in this respect, since the work of Sage *et al.* (1934) gives no information below about 20° C. Some new measurements were therefore made in these laboratories by Mr H. C. Lu to bridge the gap between 20° C and the normal boiling-point of propane. Results of these determinations, which cover a range of pressure from 5 to 60 atm., are given in table 6.

TABLE 6. *P*-*v*-*T* RELATIONSHIPS FOR LIQUID PROPANE
(H. C. LU) (VALUES OF *v*)

Pressure	Temp. ° C					
atm.	−30·5	−20·0	−10·05	+0·25	10 2	20·5
5	1·7403	1·7801	1·8220	1·8693	—	1·9863
10	1·7351	1·7766	1·8189	1·8649	1·9179	1·9796
20	1 7299	1·7706	1·8121	1·8570	1·9083	1 9686
30	1·7271	1·7655	1·8056	1·8500	1 8993	1·9587
40	1·7243	1·7612	1 7990	1·8440	1 8912	1·9489
50	1·7215	1·7579	1·7936	1·8389	1·8836	1·9392
60	1·7187	1·7547	1·7892	1·8345	1·8775	1·9295

The method employed was to take a measured volume of the purified gas and condense it into a calibrated glass capillary, surrounded by a transparent cooling jacket whose temperature was thermostatically controlled. The thread of liquid propane was confined at either end by mercury, by means of which it was subjected to varying pressures, measured with a dead-weight gauge. The length of the thread was determined by means of a kathetometer and its volume calculated from the tube calibration. Further details of the procedure and a description of the apparatus employed will be given in a future communication.

The properties of the supercooled liquid, calculated from the results obtained and from the data of Sage, are summarized in tables 2–5. The new measurements at about 20° C show good agreement with those of Sage at the same temperature.

The entropy-temperature diagram

In the accompanying chart the main plot is of entropy on a temperature base. An auxiliary diagram has also been drawn up, based upon results summarized in the tables, giving the relationship of enthalpy to temperature. With the aid of this diagram a number of isenthalpic lines have been superimposed on the main entropy chart. From the combined P-v-T data of Dana, Sage and Beattie, several isometric lines have been added.

Thermodynamic properties at low pressures

In order to complete the work, particulars of some calculations at pressures below atmospheric will be given.

Accurate determinations of the saturated vapour pressure in this region have been carried out by Hartneck and Edse (1938) whose figures have been utilized The data of Dana and Sage on the specific heat of the saturated liquid can be correlated fairly well by the expression

$$c_s = 2{\cdot}0555 - 0{\cdot}012505T + 0{\cdot}000025628T^2, \tag{20}$$

which was employed for extrapolation to low temperatures. On substituting in equation (1) and integrating, we obtain, taking the same zero-point of entropy as before,

$$s = 4{\cdot}7330 \log_{10} T - 0{\cdot}012505T + 0{\cdot}000025628T^2 - 8{\cdot}982. \tag{21}$$

The enthalpy of the saturated liquid was calculated from equation (2), v_s being extrapolated from Dana's v_s-T data, which yields a linear relationship in this region.

TABLE 7. ENTROPY AND ENTHALPY OF THE SATURATED LIQUID AND VAPOUR AT LOW PRESSURES

Pressure atm.	Temp °C	Saturated liquid		Saturated vapour	
		h	s	h	s
0·1	−83·8	−23·42	−0·112	85·97	0·466
0·2	−73·25	−17·20	−0·080	90·47	0·458
0·3	−66·15	−12·93	−0·059	93·50	0·455
0·4	−60·9	−10·21	−0·046	95·25	0·451
0·5	−57·7	− 8·28	−0·037	96·58	0·450
0·6	−53·15	− 6·10	−0·027	97·86	0·446
0·7	−49·95	− 4·32	−0·019	98·99	0·444
0·8	−47·05	− 2·52	−0·011	100·19	0·443
0·9	−44·5	− 1·15	−0·005	101·02	0·442
1·0	−42·2	0	0	101·60	0·440

The latent heat was determined by means of equation (13), giving, with the aid of equations (3) and (4), the properties of the saturated vapour. Results in the saturated region are summarized in table 7.

P-v-T relationships for propane are not available at low pressures but it is justifiable to assume adherence to the ideal gas laws. This implies the assumption

$$\left(\frac{\partial c_P}{\partial P}\right)_T = 0, \tag{22}$$

and the equation (15) becomes applicable to all pressures below atmospheric. The expression (16) can therefore be employed in the form

$$s_P = 0{\cdot}4802 \log_{10} T + 0{\cdot}000668T + K, \tag{23}$$

and by solving for the saturated state, the values of K shown in table 8 are obtained. The entropy of the superheated vapour at any temperature and pressure can therefore be obtained by working out the expression (23) using a value of K appropriate to the pressure concerned.

TABLE 8. VALUES OF K IN THE EQUATION (23) FOR PRESSURES BELOW 1 ATM.

Pressure atm.	Value of K	Pressure atm.	Value of K
0·1	−0·754	0·6	−0·826
0·2	−0·780	0·7	−0·833
0·3	−0·795	0·8	−0·839
0·4	−0·808	0·9	−0·846
0·5	−0·814	1·0	−0·849

For an ideal gas

$$\left(\frac{\partial h}{\partial P}\right)_T = 0, \tag{24}$$

and the equation (17) may be used directly to determine the enthalpy at any temperature, if the pressure is atmospheric or less.

The author is indebted to Mr H. C. Lu for carrying out measurements of the compressibility of liquid propane, and to the Department of Scientific and Industrial Research for a grant towards the expenses of the investigation.

LIST OF SYMBOLS AND UNITS EMPLOYED

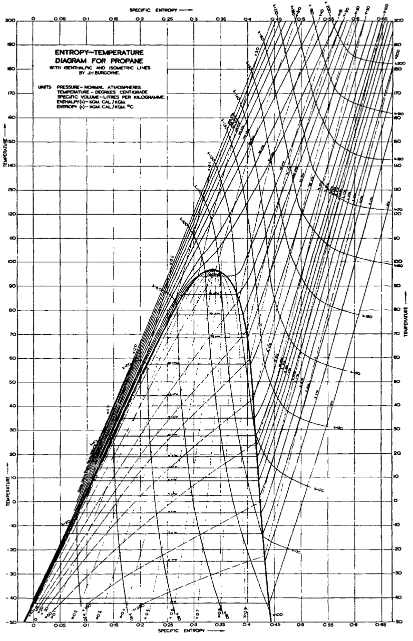
P = pressure (atmospheres).
v = volume (specific) (litres/kg.).
T = temperature (° abs.).
c = specific heat (kcal /° C).
l = latent heat of vaporization (kcal./kg.).
R = gas constant (0·00198 kcal.).
n = number of gram-molecules.
μ = Joule-Thomson coefficient (° C/atm.).
h = enthalpy (specific) (kcal /kg.).
s = entropy (specific) (kcal./kg. ° C).
$a, b, c, K, \lambda_0, \epsilon$ are constants.

Subscripts.

s refers to the saturated state.
c refers to the critical state.
P, T and h refer respectively to states of constant pressure, temperature and enthalpy.

REFERENCES

Beattie, Poffenberger and Hadlock 1935 *J. Chem. Phys.* **3**, 96.
Burrell and Robertson 1915 *J. Amer. Chem. Soc.* **37**, 2188.
Cox 1934 *Petrol. World*, **31**, 44
Dana, Jenkins, Burdick and Timm 1926 *Refrig Engr*, **12**, 387.
Francis and Robbins 1933 *J. Amer. Chem. Soc.* **55**, 4339.
Hainlen 1894 *Liebigs Ann.* **282**, 229.
Hartneck and Edse 1938 *Z. phys. Chem.* **182**, 220.
Hicks-Bruun and Bruun 1936 *J. Amer. Chem. Soc.* **58**, 810.
Justi 1936 *Z. Ver. dtsch. Ing.* **80**, 103.
Lebeau 1905 *C.R. Acad. Sci., Paris*, **140**, 1454
Maass and Wright 1921 *J. Amer. Chem. Soc.* **43**, 1098.
Meyer 1894 *Ber. dtsch. chem. Ges.* **27**, 2767.
Olszewski 1889 *Abh. Krakau. Akad. Wiss.* (see also 1894, *Ber. dtsch. chem. Ges.* **27**, 3305).
Osborne and van Dusen 1918 *Bull. U.S. Bur. Stand.* **14**, 439
Ragatz 1934 *Petrol. World*, **31**, 43.
Sage, Evans and Lacey 1939 *Industr. Engng Chem.* **31**, 763
Sage, Kennedy and Lacey 1936 *Industr. Engng Chem.* **28**, 601.
Sage and Lacey 1935 *Industr. Engng Chem.* **27**, 1484.
Sage, Nysewander and Lacey 1940 *Industr. Engng Chem.* **32**, 118.
Sage, Schaafsma and Lacey 1934 *Industr. Engng Chem.* **26**, 1218.
Sage, Webster and Lacey 1937 *Industr. Engng Chem.* **29**, 1309.


ENTROPY—TEMPERATURE DIAGRAM FOR PROPANE
WITH ISENTHALPIC AND ISOMETRIC LINES
BY J.H BURGOYNE.
UNITS PRESSURE – NORMAL ATMOSPHERES.
TEMPERATURE – DEGREES CENTIGRADE
SPECIFIC VOLUME – LITRES PER KILOGRAMME.
ENTHALPY (h) – KGM. CAL./KGM.
ENTROPY (s) – KGM. CAL./KGM. °C
SPECIFIC ENTROPY
0
0·05
0·1
0·15
0·2
0·25
0·3
0·35
0·4
0·45
0·5
0·55
0·6
0·65
TEMPERATURE
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
h=200
h=190
h=180
h=170
h=160
h=150
h=140
h=130
h=120
h=110
h=100
h=90
h=80
h=60
h=40
h=20

(Facing p. 311)

The quenching of the resonance radiation of sodium

By R. G. W. Norrish, F.R.S., and W. MacF. Smith
Laboratory of Physical Chemistry, Cambridge

(*Received* 17 *June* 1940)

The effective cross-section for quenching of sodium resonance radiation by the saturated and unsaturated hydrocarbons, tertiary amines and several diatomic molecules has been measured. The results indicate that with regard to quenching ability the gases fall into two groups, one comprising the saturated hydrocarbons and the inert gases and the other the unsaturated hydrocarbons and the amines. The difference in behaviour is too pronounced to be explained in terms of the discrepancy between the amount of energy the sodium atom gives up and the quenching molecule can receive, but may be attributed to the presence of unsaturation in the molecule. Within any one series the number of atoms in the molecule apart from the unsaturated centre seems to have little influence on the quenching ability, and it has been concluded that the quenching ability may be regarded as proceeding from a centre of unsaturation. The results have been qualitatively considered in the light of Stearn and Eyring's theory of non-adiabatic reactions, and in the terminology of the theory of the intermediate complex we may say that the presence of unsaturation manifests itself in a relatively large transmission coefficient

The fate of electronic energy in photochemical processes is one of the principal problems in the elucidation of the mechanism of these reactions. In the most general case of the photolysis of a polyatomic molecule, the energy of excitation may be re-emitted as fluorescence, may give rise to unimolecular decomposition or be degraded into thermal energy by processes independent of collision, while other processes of reaction or degradation involving collisions may also occur. The simplest way in which these processes of degradation may be studied, dissociated from the complications of chemical reaction, is by the study of the quenching of the resonance radiation of a monatomic gas, for in such systems there is but one process which is independent of collision, namely, the re-emission of the excitation energy as resonance radiation, while by a suitable choice of the quenching substance we can limit at will the processes dependent upon collision to degradation, of the electronic energy or to chemical reaction.

Work of this type has been done on the quenching of the resonance radiation of several atomic vapours particularly that of mercury (Stuart 1925; Noyes 1927; Mitchell 1928; Bates 1928, 1930, 1932; Evans 1934) and sodium (Mannkopff 1926; Winans 1930; Terenin and Prileshajewa 1931, 1932; Kisilbasch, Kondratjew and Leipunsky 1932). The data obtained for sodium are restricted to a few unrelated compounds and those for mercury suffer because there is the possibility of some ambiguity in their interpretation. Quenching of a mercury atom in the first excited level $^3P_1$ probably occurs more often through transfer to the near metastable $^3P_0$ state rather than to the ground level $^1S_0$ In contrast with mercury the first excited level of the sodium atom is not associated with any metastable state, and quenching must of necessity result in a transition direct to the ground state. The magnitude of the energy change involved is so large that we may expect to find quite different quenching relationships in these processes and possibly some indications of the factors in molecular structure which most influence the degradation of electronic energy to heat.

The effective cross-section mentioned above is a convenient quantitative measure of the quenching efficiency of a gas. It is defined as follows. The number of collisions per sec. per c.c. per excited atom is, according to kinetic theory, given by the expression

$$2n\sigma^2\sqrt{\{2\pi RT(1/M_1+1/M_2)\}} = Z_Q,$$

where n is the number of molecules of foreign gas per c.c. and M_1 is the molecular weight of the atomic vapour and M_2 the molecular weight of the foreign gas, and σ is the maximum distance between centres at which collision occurs. The value of σ^2 which must be used in this relation to give the number of collisions per excited atom per c.c. which is required to account for the observed quenching, assuming every collision to be efficient, is called the effective cross-section.

The quantity measured experimentally in investigations on quenching is the quenching ratio Q which is the ratio of the amount of radiation emitted in the presence of foreign gas to that emitted in the absence of it. Z_Q, and consequently σ^2, may be simply related to the quenching ratio if it be assumed that the presence of foreign gas does not affect the amount of radiation absorbed and if the radiation emitted is proportional to the concentration of excited atoms.

These conditions hold if the vapour pressure of the absorbing atoms is so low that only primary resonance radiation is emitted and if the effect of Lorentz broadening on the absorption line is negligible and if there is no appreciable absorption or emission stimulated by collision.

The ratio of the concentration of excited atoms in the presence of foreign gas to that in its absence is controlled completely, under these circumstances, by the probability of deactivation relative to that of emission of radiation in unit time. This ratio is equal to the quenching ratio. In terms of the collision frequency Z_Q and the lifetime τ of an excited atom, the relation is

$$Q = 1/(1+\tau Z_Q).$$

This is known as the Stern-Volmer relation. Since Z_Q is proportional to the pressure of added gas a linear relationship between $1/Q$ and the pressure is indicated. If the above expression for Z_Q is substituted in this relation there is obtained an expression for σ^2 in terms of the quenching ratio, the concentration of quenching molecules and the lifetime of an excited atom.

The process of quenching is a collision of the second kind and as such has been considered from a quantum mechanical standpoint (Kallmann and London 1929; Morse and Stückelberg 1931). Calculations based upon hypothetical curves representing the potential energy of the interacting particles have shown that the effective cross-section for the process depends on the nature of the interaction between colliding particles, on the relative velocity before collision, and on the change of relative kinetic energy on collision, that is, on the difference between the amount of energy the excited atom has to give and the quenching molecule to receive. Experimental data have given some support to the assumption that the effective cross-section increases with decrease of the discrepancy between the differences of energy level in the colliding particles. We would expect such a correlation to have general application only if the nature of the potential field of the interacting particles were related in a unique way to the value of differences of energy level in the isolated particles. If no simple relation holds we may expect that the effective cross-section will increase in regular manner with decrease in the energy discrepancy only in a particular series where the law of interaction of the particles is similar.

It is possible to apply Stearn's and Eyring's treatment of non-adiabatic reactions (Stearn and Eyring 1935) to the process of quenching. The process may be visualized as involving a transition between two energy surfaces, one representing the potential energy for configurations of the system composed of excited atom and quenching molecule, and the other for the system composed of normal atom and quenching molecule. This treatment emphasizes that three factors are dominant in controlling the rate constant for the process, an entropy of activation, an energy of activation and a quantity which is the probability that a system in the activated state yields the products normal atom and molecule. This last quantity is a function of the slopes of the surfaces at the region of closest contact and of their energy difference and depends upon the velocity of the point representing the states of the interacting particles. In the absence of knowledge concerning the intermediate states for particular sets of molecule and atom this method can only be used to consider the results from a qualitative standpoint. This treatment does, however, provide an illuminating method of classifying the factors which control quenching efficiency, and will be further considered later in this paper.

It has been suggested that the effective cross-section for quenching is dependent on chemical forces of interaction which may be estimated by the heat of reaction of the reaction $Na + AB \rightarrow Na\,A + B$ (Kondratjew and Siskin 1936). It appears that the quenching process and non-adiabatic reactions may be treated in almost identical fashion, but these two processes are quite distinct, and the potential surfaces for describing the processes are quite different. Even if the energy change in such a hypothetical reaction could be estimated it is difficult to see on what grounds it would be expected to be a quantitative measure of the effect of the factors which control quenching.

The main object of this work is to adduce more experimental evidence than is at present available as to the behaviour of different molecules in quenching the resonance radiation of sodium. Therefore we have used different homologous series attempting to assess in each case the specific qualities of molecular structure which can give rise to quenching.

We have found that with the saturated hydrocarbons quenching is slight, and is not markedly dependent upon the number of degrees of freedom. On the other hand, quenching is marked with molecules which possess a high degree of unsaturation, either in the form of double bonds or lone pairs of electrons. Apparently the magnitude of the energy discrepancy is not the main factor controlling quenching efficiency.

EXPERIMENTAL

Apparatus

The general experimental arrangement is indicated in figure 1. The sodium lamp *L*, used as source, was a G.E.C. 'Osira' laboratory model. It was mounted in a metal box from which the radiation could escape via lens *A* to the quenching cell *C* and via tube *R* and mirror *Q* to the opalescent screen *S*. The intensity of the illumination at the screen could be controlled by the iris *I*. The quenching cell had the form indicated and had a blackened light trap *T* opposite the observation window. This cell was placed in a double-walled metal box through which was circulated hot air. The air was heated by blowing it through a tube heated electrically, and its temperature was controlled manually. The temperature was observed by means of a thermometer, graduated in tenths of a degree, and could be kept constant to better than 0·2° C. The exciting beam entered the thermostat through the two glass windows M and the resonance radiation could be observed through two others N. The opalescent screen S served as comparison source with which the intensity of the resonance radiation could be compared with the aid of the photometer P. (The two sources were about 3·2 in. from the base of the instrument and effectively 2 in. apart.) Images of the two sources were formed in the eye aperture of the instrument, and the brightness of the two halves of the field were matched by means of a wedge adjustment W placed in the path of the radiation from the comparison source. The ratio of the two brightnesses B_1 and B_2 could be obtained from the two scale readings P_1 and P_2 with the relation $\log B_1/B_2 = \alpha(P_1 - P_2)$, where α is a constant for the instrument and had a value of 0·077 for light of wave-length 5890 A. The ratio of the readings obtained in the presence and in the absence of gas is the quenching ratio. Minor variations in the intensity of the resonance radiation due to variations in the lamp intensity

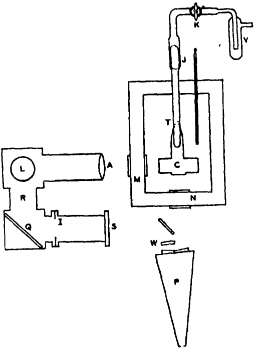


FIGURE 1. Apparatus.

were accompanied by corresponding variations in that of the comparison source and so did not affect determination of the quenching ratio. The photometer was built and calibrated by Adam Hilger, Ltd.

The tubing exterior to the thermostat was wound with nichrome wire up to the magnetically operated stopper *J*. Gas was admitted to the cell via this stopper and the stopcock *K*, from the vessel *V*. This vessel was connected with a glass Bourdon gauge and to bulbs for gas storage. The pressure in the quenching vessel could be determined by direct reading with stoppers *J* and *K* open Experiments showed that the use of stopper *J* and the heating of the tubing exterior to the thermostat was not necessary.

Sodium was purified and introduced into the quenching vessel by the method suggested by Dunoyer (1912). The sodium was heated *in vacuo* for about 40 min. to a temperature at which sodium was distilling at an appreciable rate so that all oil enclosed in the metal might be removed. The bulb containing the sodium was then broken off from the vacuum line and was sealed into the end-bulb of a distillation chain It was difficult to distil the sodium from this bulb because of the two glass walls, and the end bulb was therefore inclined so that the sodium on melting could run into the second bulb from which distillation was easy after sealing off the end one. Distillation and sealing off was continued until sodium was introduced into the tubing connected directly to the quenching vessel. During the distillation, the distillation chain was directly connected with the pump to remove any water or carbon dioxide that was introduced during the initial sealing in of the sodium Sodium was then driven from the tubing into the quenching vessel, leaving however a mirror of sodium lining the walls of the connecting tubing.

Purification of materials

Helium was purified by passage through charcoal cooled to liquid-air temperature.

Methane was prepared by dropping methyl iodide dissolved in alcohol on to a copper-zinc couple. The gas was passed through a 20 % solution of fuming sulphuric acid, and was stored over a solution of sodium hydroxide and dried and purified by distillation *in vacuo*.

Ethane was prepared by the hydrolysis of zinc diethyl. The tube of zinc diethyl was broken *in vacuo* by a magnetic hammer, and water from which the air had been completely removed was allowed to drop on it. The gas evolved was passed over phosphorus pentoxide and fractionated before use.

Cyclohexane was a B.D.H. product specially purified for spectroscopy.

Iso-octane was a fraction of iso-octane obtained from 'Dr Schuchardt of Görlitz'.

Methyl cyclohexane was a product of Howard and Sons.

Propane and butane were cylinder gases fractionated several times before use.

Nitrogen was prepared by heating sodium azide *in vacuo*. It was first heated while connexion to the vacuum pump was maintained until it began to decompose. The quantity of sodium azide used during a preparation was small, as decomposition occurred with some violence.

Ethylene was prepared by dropping ethyl alcohol on to syrupy phosphoric acid which had been dehydrated and heated to 210° C. The preparation was carried out *in vacuo* and the gas passed through a trap cooled with solid carbon dioxide and ether and then over phosphorus pentoxide. The gas was fractionated from liquid air before it was used.

Propylene and butylene were cylinder gases fractionated *in vacuo* several times before use.

Pyridine was distilled, treated with KOH and then with sodium.

Trimethyl amine was prepared from the hydrochloride by treatment with an aqueous solution of potassium hydroxide. The hydrochloride was purified by crystallization from alcohol, and the free base was dried over potassium hydroxide and was fractionated three times.

The ethers were both fractionally distilled and treated with sodium prior to use.

1-5-Hexadiene was fractionally distilled before use.

Cyclohexene was a B.D.H. product and was fractionally distilled before use.

Carbon monoxide was prepared by dropping concentrated sulphuric acid on sodium formate and was purified by passage through solid carbon dioxide and ether, over soda lime and caustic potash and distillation *in vacuo*.

RESULTS

Some difficulty was encountered in the initial experiments with a fading in the intensity of the radiation emitted from the quenching vessel after addition of a quenching gas. It was found however, that with an ample supply of sodium in the vessel this no longer occurred, and the relation between the quenching ratio and the pressure of quenching gas was the same on stepwise removal as it was upon stepwise admission.

Tables recording the quenching measurements have been deposited in the archives of the Royal Society; the results can be inferred from figures

2 and 3. The gases have, for convenience, been divided into two groups strongly and weakly quenching; it will be noted that experimentally the two groups are quite distinct. In figure 2, $1/Q$ is plotted against pressure for the strongly quenching gases and in figure 3, Q is plotted against pressure for the weakly quenching gases.

In table 1 are given the values of the effective cross-section for the strongly quenching gases for the weakly quenching gases. The calculations for the strongly quenching gases were made with values of the pressure necessary to bring about a diminution in the intensity of the radiation emitted to 0·50 or 0·67 of that emitted in the absence of gas. As pointed out below the quenching ratio for this group of gases is controlled almost entirely by the quenching process, and the values of the effective cross-sections calculated from these measurements are quantitative measures of the quenching efficiency. The calculations for the weakly quenching gases were made with values of the pressure necessary to bring about a value of the quenching ratio of 0·97. The choice of a value of the quenching ratio corresponding to low pressures was made because, as will

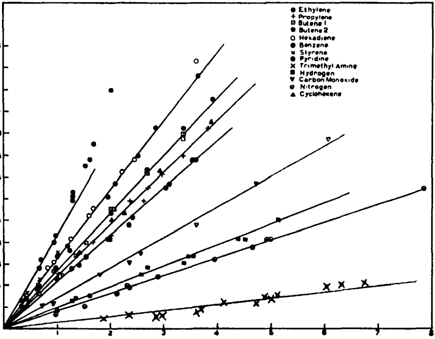


FIGURE 2

be indicated, factors other than quenching exert an appreciable influence on the ratio at higher pressures. The values used in the calculations were estimated from figure 3 and obviously cannot be expected to yield quantitative estimates of the cross-section, but they should indicate the order of quenching efficiency.

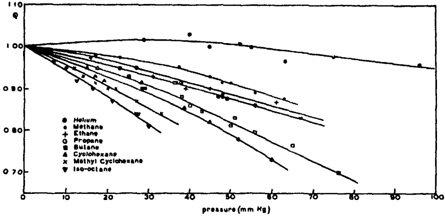


FIGURE 3

TABLE 1. QUENCHING DIAMETERS

Quenching gas		Effective cross-section (σ^2) (cm.^{2} × 10^{16})
Strongly quenching gases:	Ethylene	44·0
	Propylene	52·0
	Butylene (1 and 2)	58·0
	Cyclohexene	59·0
	Nitrogen	14·5
	Hydrogen	7·4
	Carbon monoxide	28·0
	Benzene	75·0
	Hexadiene	76·0
	Styrene	78·0
	Pyridine	104·0
	Trimethyl amine	6·5
Weakly quenching gases:	Methane	0·11
	Ethane	0·17
	Propane	0·2
	Butane	0·3
	Cyclohexane	0·4
	Methyl cyclohexane	0·55
	Iso-octane	0·8

To the weakly quenching gases belong the saturated hydrocarbons and helium as well as other inert gases (Von Hamos 1932). It will be noted that the effective cross-section for the saturated hydrocarbons is in all cases much less than that holding in ordinary kinetic processes and increases somewhat with increase in complexity in the quenching molecule.

The unsaturated hydrocarbons belong to the strongly quenching group. Those with one double bond have an effective cross-section greater than the normal collision diameter, and this marked quenching ability increases slightly with increase in the length of the carbon chain but as indicated by the results obtained with the two butylenes, butene 1 and butene 2, is independent of the position of the double bond in the chain. Hexadiene with two unsaturated linkages possesses an effective cross-section considerably greater than that of those compounds with only one. It is interesting to note that benzene, hexadiene and styrene have approximately the same quenching diameter.

Pyridine and trimethyl amine show very marked differences in quenching ability. Nitrogen, hydrogen and carbon monoxide have effective cross-sections not far from the values indicated by the kinetic theory.

The results observed with styrene deserve special mention. After admission of styrene to the quenching vessel, the pressure of styrene decreased gradually to zero. The amount of quenching decreased with decrease in pressure and was equal in magnitude to that obtained when styrene was freshly introduced to that particular pressure. In view of the fact that styrene polymerizes slowly in the liquid phase at this temperature and that sodium does increase the rate of polymerization of other unsaturated hydrocarbons (Schulz and Husemann 1936, 1937) such a result is not unexpected.

Experiments were also carried out with butyraldehyde, diethyl ketone and diethyl selenide, but in all cases chemical reaction made the quenching measurements impossible or uncertain. Butyraldehyde and diethyl ketone were removed immediately by reaction with sodium. When diethyl selenide was admitted to the quenching cell a very pronounced diminution in the intensity of the radiation occurred. The intensity then rose reaching a maximum in about 2 min, after which it gradually decreased, finally reaching a constant value. During this interval the pressure in the quenching vessel increased to over twice the original.

DISCUSSION OF EXPERIMENTAL METHOD

As pointed out above, the experimental results may be simply interpreted only in the absence of imprisonment of radiation and of Lorentz broadening of the absorption line.

At a temperature of 150° C the resonance radiation appeared to come only from the path of the exciting beam, and as in this work the temperature of the sodium vapour was close to 130° C, at which temperature the vapour pressure is considerably less than at 150°, radiation imprisonment will have a negligible effect on these measurements.

In considering the effect of Lorentz broadening the division of gases used into the two groups strongly quenching and weakly quenching is convenient. For the strongly quenching gases collision diameters were calculated from measurements for which the gas pressure was less than 2 mm. of mercury. With the weakly quenching gases measurements were carried out at pressures up to 70 mm.

From the measurements on Lorentz broadening of the absorption line of sodium (Schütz 1927), we may estimate that, at 130° C and with a pressure of helium of 5 mm. of mercury, the Lorentz breadth is less than 3% of the Doppler breadth. As the effective cross-section for Lorentz broadening does not vary markedly from gas to gas we may assume that for the strongly quenching gases the effect of Lorentz broadening will be small compared with that of quenching. However, with the weakly quenching gases Lorentz broadening may have a considerable influence on the observed value of the quenching ratio, and the quenching curves for the weakly quenching gases indicate a deviation from linearity in the relation between $1/Q$ and pressure. The quenching curve for helium indicates that at pressures of less than 40 mm there is an increase in the intensity of the radiation above that emitted in the absence of gas. This may be the result of an increase in the absorption of radiation due either to Lorentz broadening of the absorption line or to collision-stimulated absorption and emission. It is impossible to estimate the contribution of these two effects, but the slope of the quenching curve for helium suggests that their net influence is not great at pressures less than 5 mm At the higher pressures the effect of Lorentz broadening and shift may be to diminish the amount of radiation absorbed. This would seem to be the case with the saturated hydrocarbons whose quenching curves indicate a pronounced decrease in the radiation emitted at higher pressures

Reaction of added gas with unexcited sodium atoms might cause a reduction in the intensity of the radiation emitted. As the pressure of gas is increased the rate of removal of sodium atoms by reaction is increased and their rate of supply by diffusion into the region considered is lessened. Consequently there will be a diminution in the concentration of sodium and in the amount of radiation absorbed from the exciting beam and consequently in the concentration of excited atoms. The drop in intensity will then be greater than it would be in the absence of reaction. Such behaviour is shown by the ethers (figure 4), the quenching curves of which indicate quenching at higher pressures far in excess of that to be expected from the Stern-Volmer relation and the low pressure values. A pronounced dependence of the quenching on the temperature is also shown. With increase in temperature the amount of quenching decreases markedly. Apparently the increased rate of supply of sodium atoms into the region observed more than compensates for any increase in the rate of removal by reaction at the higher temperature as far as any effect on the equilibrium concentration of the sodium is concerned. This behaviour may be contrasted with that of nitrogen for which the quenching curves remain linear and the amount of quenching increases slightly with increase of temperature.

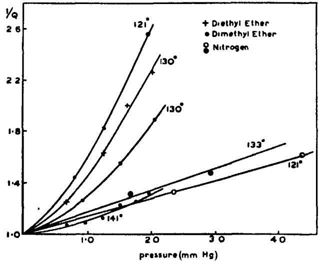


FIGURE 4

A dependence of the sodium atom concentration on the rate of diffusion of the sodium atoms might possibly give rise to local differences in the concentration of sodium atoms in the presence of foreign gas and consequently to different quenching curves for observations from different sections along the illuminating beam. Quenching measurements carried out with nitrogen at different regions along the exciting beam indicated no appreciable variation in the quenching curves. Considered along with the fact that the quenching curves in which $1/Q$ is plotted against pressure are linear and reproducible for the strongly quenching gases with the exception of pyridine, this evidence, while negative and particular, supports the suggestion that the removal of sodium atoms by reaction has a negligible effect on the quenching ratio. Pyridine when treated with sodium prior to use, yielded a reddish deposit and consequently it is not surprising that reaction with unexcited sodium atoms occurs at a rate sufficient to affect the quenching ratio.

With weakly quenching gases quite high pressures were used, and it is possible that the bend in the quenching curves is partly due to reactions of sodium vapour with a trace of impurity.

We may sum up these considerations by saying that the results for the strongly quenching gases may be taken to indicate quantitatively the quenching efficiency, but the results for the weakly quenching gases can give only semi-quantitative indications.

DISCUSSION

The contrast in the behaviour of the hydrocarbons in the quenching of the resonance radiation of sodium and that of mercury is evident from table 2, where the data obtained by Bates (1930, 1932) for mercury is listed along with ours for sodium. For the quenching of mercury resonance radiation a marked dependence of the effective cross-section of the saturated

TABLE 2. QUENCHING DIAMETERS

Quenching gas		Effective cross-section (σ^2)	
		(cm.^{2} × 10^{16}) (Mercury)	(cm.^{2} × 10^{16}) (Sodium)
Saturated:	Inert gases	~ zero	~ zero
	Methane	0·059	0 11
	Ethane	0·421	0·17
	Propane	1·60	0·2
	Butane	4 06	0·3
	n-Heptane	24 0	—
	223-Trimethyl butane	19 7	—
	Cyclohexane	—	0·4
	Methyl cyclohexane	—	0·55
	Iso-octane	—	0 8
Unsaturated:	Ethylene	—	44·0
	Propylene	—	52·0
	Butene 1	—	58·0
	Butene 2	—	58·0
	Heptene 1	45·8	—
	Heptene 3	55 9	—
	Benzene	41·9	75 0

hydrocarbons is shown varying from a very small value for methane to one greater than the kinetic value for heptane. The results for sodium, on the other hand, indicate a small fairly constant effective cross-section for quenching by all the saturated hydrocarbons. Zemansky (1930) has considered that in the quenching process excited mercury is lowered to the nearby metastable $^3P_0$ state by all the hydrocarbons, and with this assumption has used the data of Bates to examine the relation between the effective cross-section and the discrepancy between the amount of energy given up by the mercury atom and that which can be received by the hydrocarbon molecule. When the effective cross-section is plotted against the energy discrepancy a smooth curve can be drawn through the points, and this curve is also satisfactory for molecules such as nitric oxide, carbon monoxide and ammonia which possess appreciable dipole moments. Bates has pointed out that *n*-heptane did not fit into the curve, and that in view of the photosensitized reaction which occurs with all the hydrocarbons except methane the metastable state is probably not the one attained in quenching. Norrish (1939), on the other hand, has correlated the quenching of mercury resonance radiation by saturated hydrocarbons with the number of degrees of freedom of the quenching molecule. The figures of Bates shown above indicate a remarkable increase as we pass up the homologous series of paraffins and it seems probable that when the energy transfer involved in quenching is small, and of the order of magnitude of vibrational transitions in the quenching molecule, as it is with mercury (in the $3P_1 \rightarrow 3P_0$ transition), then in the absence of other influences such as unsaturation, the probability of quenching is largely determined by the number of vibrational degrees of freedom.

In the quenching of sodium resonance radiation by hydrocarbons there is no possibility of photosensitized reaction and there is only one transition possible and that involves a change in energy in the sodium atom of 2·094 eV compared with the change of only 0·218 eV associated with the transition of the mercury atom to the metastable state. The pronounced difference in the behaviour of the saturated hydrocarbons with excited sodium and excited mercury may be due to this difference in the magnitude of the energy transferred, for it far exceeds the vibration levels of the quenching molecules.

The data for sodium indicate that we may divide the substances used into two groups, the first comprising the saturated hydrocarbons and the inert gases and possessing small cross-sections for quenching, and the second comprising unsaturated compounds—hydrocarbons with double linkages and amines with unpaired electrons—and possessing cross-sections for quenching near the ordinary kinetic values. The various values of the energy discrepancy for the saturated hydrocarbons cannot be sufficiently different from those of the unsaturated hydrocarbons to explain this pronounced difference in quenching efficiency. It may be noted that the molecules of the first group are inert and possess small external fields, while those of the second are unsaturated and possess large external fields. Apparently the sphere of 'quenching influence' is somehow related to the magnitude of the external field. This is in contradiction to Zemansky's conclusion that the nature of the external field has little influence on the cross-section for the quenching process, but is not inconsistent with the theoretical work on collisions of the second kind (Kallmann and London 1929; Morse and Stückelberg 1931), where it is shown that the form of the quenching curve depends on the nature of the interaction between the particles concerned.

On the border of the second group may be placed hydrogen, nitrogen and carbon monoxide. Nitrogen and carbon monoxide may be regarded as possessing a potential unsaturation which will be made operative by the proximity of a sodium atom. Along with hydrogen they probably suffer greater distortion on the approach of a sodium atom than any other molecule of the group, as both the hydride and azide are easily formed. While this interaction brings about the conditions necessary for the transition process it is considerably less effective than the unsaturation conferred by the presence of a double carbon to carbon linkage.

The quenching process may be pictured as involving a transition of the representative point for the system, sodium atom and quenching molecule, from one potential surface to another. Eyring's expression for the specific rate constant for quenching as developed from such a picture is

$$K = \kappa(kT/h)\exp(\Delta S/R)\exp({}^{-\Delta E}/RT),$$

where ΔS and ΔE are the entropy and energy of activation and κ is the transmission coefficient. The entropy and energy of activation give quantitative expression to the factors which control the concentration of the quasi-molecule composed of excited atom and quenching molecule in the states from which transition to the lower surface may occur, and the transmission coefficient represents the probability that, once the representative point is in the region of closest approach of the surfaces, transition to the lower surface and motion along it corresponding to separation of the atom and molecule occurs.

The influence of the energy and entropy of activation on the quenching efficiency is examined below, where it is pointed out that while it may be appreciable, it cannot be sufficient to explain the great difference in quenching ability shown by the two groups of quenching gases. The *transition probability*, however, depending upon the nature of the potential surfaces, may vary enormously and is doubtless the factor to which pronounced differences in quenching ability may be attributed.

The large cross-section for quenching by the unsaturated compounds suggests that transition occurs at such a separation of atom and molecule that the repulsive force, and consequently the energy of activation, is small. For the saturated hydrocarbons there is the possibility that the potential surfaces approach only at distances so small that considerable repulsive forces are called into play, and so that there is an appreciable energy of activation. Measurements made with the saturated hydrocarbons at other temperatures indicated no marked change in the quenching efficiency, as would be expected if there were a large energy of activation. The experimental evidence is not sufficient however, to draw from it any definite conclusions about the energy of activation.

An appreciable positive entropy factor indicates that there is greater randomness in the intermediate complex than in the normal atom and molecule and an appreciable negative factor indicates that the state for the intermediate complex demands some peculiar orientation of the molecule and the atom. (There must always, of course, be the factor corresponding to the condition that the atom and molecule are in each other's neighbourhood in the intermediate state.) Quenching, in general, may be expected to involve less distortion in the participating molecules than does chemical reaction, and consequently no great alteration in the randomness of the intermediate state compared with that of the separated atom and molecule. Likewise, although it is conceivable that the molecule might have to attain a peculiar orientation of its own atoms for quenching, there is no reason to suppose that this would be necessary for the substances employed in this work. On the other hand, a particular orientation of molecule relative to the excited atom may be necessary. The large effective cross-section for the unsaturated hydrocarbons, and the fact that its magnitude is nearly independent of the position of the double bond in the molecule, suggests that the necessity for orientation does not play an important part in this quenching process. The small effective cross-section of tri-methyl amine is consistent with the assumption that a peculiar orientation of the amine molecule relative to the excited atom is necessary. Pyridine, on the other hand, has a very large effective cross-section and may be classified with the compounds possessing double linkages and for which orientation is not an important prerequisite for transition.

A small variation in the effective cross-section in any series of compounds is to be expected on theoretical grounds, and is attributable to variation in the transition probability. Other factors remaining unaltered, the probability of transition depends on the velocity of the representative point in the region of the potential surface corresponding to the intermediate state, and in the quenching process this runs parallel to the relative velocity of the sodium atom and quenching molecule, for with increase in molecular weight the average velocity of the molecule will decrease and the effective cross-section increase. In view of the small increase actually observed with the hydrocarbons with one unsaturated linkage it appears that the number of atoms in the molecule has little effect on the quenching efficiency. Indeed, we may think of the ability to quench as proceeding from the unsaturated centre just as we think of the ability of a large molecule to absorb radiation as belonging to a chromophoric group. Further evidence for such a view is furnished by the two butylenes, butene 1 and butene 2, which have the same effective cross-section for quenching. It is also interesting to note, in connexion with the sphere of 'quenching influence' surrounding a molecule, that hexadiene with two unsaturated linkages has nearly as great an effective cross-section as benzene and styrene.

Acknowledgement is due to the Commissioners of the Exhibition of 1851 for the award of an Exhibition to one of us (W. MacF. S), and to the Royal Society for a grant for apparatus.

REFERENCES

Bates 1928 *Proc. Nat. Acad. Sci*, *Wash.*, **14**, 849.
— 1930 *J. Amer. Chem. Soc.* **52**, 3825.
— 1932 *J. Amer. Chem. Soc.* **54**, 569.
Dunoyer 1912 *Radium, Paris*, **9**, 177.
Evans 1934 *J. Chem. Phys.* **2**, 445.
Kallmann and London 1929 *Z. Phys. Chem.* B, **2**, 207.
Kisilbasch, Kondratjew and Leipunsky 1932 *Phys. Z. Sowjet.* **2**, 201.
Kondratjew and Siskin 1936 *Phys. Z. Sowjet.* **8**, 644.
Mannkopff 1926 *Z. Phys.* **36**, 315.
Mitchell 1928 *Z. Phys.* **49**, 228.
Morse and Stuckelberg 1931 *Ann. Phys., Lpz.*, **9**, 579.
Norrish 1939 *Trans. Faraday Soc.* **35**, 22.
Noyes 1927 *J. Amer. Chem. Soc.* **49**, 3100.
Schulz and Husemann 1936 *Z. Phys. Chem.* B, **34**, 187.
— — 1937 *Z. Phys. Chem.* B, **36**, 184.

Schütz 1927 *Z. Phys.* **45**, 30.
Stearn and Eyring 1935 *J. Chem. Phys.* **3**, 778.
Stuart 1925 *Z. Phys.* **32**, 262.
Terenin and Prileshajewa 1931 *Z. Phys. Chem.* B, **13**, 72.
— — 1932 *Phys. J. Sowjet.* **2**, 337.
Von Hamos 1932 *Z. Phys.* **74**, 379.
Winans 1930 *Z. Phys.* **60**, 631.
Zemansky 1930 *Phys. Rev.* **36**, 919.

On maintained convective motion in a fluid heated from below

BY ANNE PELLEW AND R. V. SOUTHWELL, F.R.S.

(*Received* 1 *July* 1940)

This paper examines the stability in viscous liquid of a steady regime in which the temperature decreases with uniform gradient between a lower horizontal surface which is heated and an upper horizontal surface which is cooled. The problem has been treated both experimentally and theoretically by Bénard, Brunt, Jeffreys, Low and Rayleigh, and it is known that instability will occur at some critical value of $gh^3\Delta\rho/\rho k\nu$, h denoting the thickness of the fluid layer, $\Delta\rho/\rho$ the fractional excess of density in the fluid at the top as compared with the fluid at the bottom surface, k the diffusivity and ν the kinematic viscosity. The critical value depends upon the conditions at the top and bottom surfaces, which may be either 'free' or constrained by rigid conducting surfaces.

The theoretical problem is solved here under three distinct boundary conditions, and greater generality than before is maintained in regard to the 'cell pattern' which occurs in plan. In addition an approximate method is described and illustrated, depending on a stationary property akin to that of which Lord Rayleigh made wide application in vibration theory.

Within the assumptions of the approximate theory (i.e. with neglect of terms of the second order in respect of the velocities) a particular size is associated with every shape of cell (such that 'a^2' takes a preferred value), but no particular shape is more likely than another to occur in a layer of indefinite extent (§ 31). The explanation of the apparent preference for a hexagonal cell pattern (§ 5) must presumably be sought in a theory which takes account of second-order terms. This conjecture if correct goes some way towards explaining the rather indefinite nature of observed cell-formations (cf. Low 1930, figure 10).

THE PHYSICAL PROBLEM

1. When in a stationary fluid some layer has greater density than others which lie below it, its equilibrium is plainly unstable in the sense that even

a small disturbance may result in a completely changed regime. The difference in density may either be intrinsic, as when water condenses on the upper surface of a layer of oil, or it may be due to such causes as unequal salinity or temperature: in that event it will tend in time to be counteracted by diffusion, and convective (i e. molar) motion of the fluid will not result if the diffusivity be sufficiently high. But diffusion may in turn be counteracted by some cause tending to maintain the inequality of density,—as when a fluid has its upper surface kept at constant temperature and heat is applied to it from below.* The question of stability in these circumstances has physical importance from several aspects. It appears to have been first studied by Lord Rayleigh (1916) in relation to experiments made by Bénard (1900).

In a theoretical treatment of the problem it is natural to assume that the top and bottom surfaces of the fluid are plane and horizontal so that its depth has a constant value h, and to take as the other boundary a vertical cylinder having any shape of horizontal cross-section. This cylindrical surface, or 'cell wall', may either be a material boundary (e.g. in experimental work) or it may be a surface of symmetry between adjacent cells of a 'convection pattern' occurring spontaneously in a fluid of infinite horizontal extent. In both instances we may impose the condition that *no heat is transmitted through the cell wall*, either because its material is thermally non-conducting† or in consequence of the predicated symmetry.

2. At the top or bottom surface, on the other hand, either of two conditions may be postulated· the material in contact with the fluid may be (thermally) either an insulator or conductor. In the first event no heat is transmitted, so the normal gradient of the temperature vanishes at the horizontal surface; in the second we may postulate that a uniform and specified temperature is maintained at every point. Here we shall impose the second condition at both surfaces, on the ground that only so can steady convective motion be *maintained*. Clearly, if one surface is a thermal insulator then all parts of the fluid must come in time (by conduction) to the temperature of the other, and a like regime of uniformly distributed temperature (manifestly stable) must result when both surfaces are non-conducting; so unless the temperature at each surface is kept constant at some cost in heat transmission, any instability discovered in analysis must

---

* The heat must of course be supplied from above if the fluid has a negative coefficient of thermal expansion (e.g. water close to its freezing point). Cf. § 20.

† Non-conducting material would almost certainly be used in experimental work. Heat transfer through the cell wall would entail prohibitive complication.

relate to some temporary distribution of temperature, and will be succeeded by some motion having (ultimately) the nature of a damped oscillation or subsidence.*

3. Not only the temperature but also the velocities have to satisfy appropriate conditions at the boundaries. When the fluid has viscosity, material boundaries (over which it cannot slip) will require all three components of velocity to vanish, surfaces of symmetry will entail a like number of conditions.† Detailed investigation shows (§ 11) that the problem can be formulated in terms either of the temperature (θ)‡ or of the vertical component of velocity (w), that both quantities are governed by equations of a like mathematical form, and that at every point of every boundary three conditions (in all) are imposed. Having determined θ we can deduce w and vice versa, and when w has been determined a supplementary calculation leads to the other two components of velocity, one further governing equation being involved together with one new condition at every boundary.

REVIEW OF PREVIOUS MATHEMATICAL INVESTIGATIONS

4. Rayleigh's mathematical treatment proceeds primarily in terms of w, that of Jeffreys (1926) primarily in terms of the temperature (for which he employs the symbol V). Both assume that equations expressing the conditions of neutral stability are obtained when second-order terms are neglected and all time variations made zero, i.e. when the assumed regime entails steady velocity and temperature at any one point, the velocity being everywhere small Jeffreys (§ 1) examines this assumption in some detail, showing that in effect it disregards the possibility that two exponential time-factors may coalesce and become conjugate complex quantities, so as to have finite imaginary parts when their real part vanishes. He cites Rayleigh as having disposed of this possibility in relation to boundary conditions of a special type, and himself makes appeal to experiment to confirm that it is not the type of motion that does in fact arise when equilibrium breaks down. In this paper (which follows Rayleigh in dealing primarily with w) a more detailed investigation (§§ 16–18) seems to establish

---

* It would seem that one of Jeffreys's results (case (2) of § 6) must be interpreted in this sense. Rayleigh's discussion (like this paper) is restricted to the case of two conducting surfaces.

† In the classical experiments of Bénard (cf. § 1) the bottom (heated) surface was a metallic plate, the upper (cooled) surface was usually free.

‡ More precisely, θ stands for the change of temperature due to convectional motion.

positively that ***any oscillatory motion must of necessity decay***: consequently in seeking the conditions for maintained convective motion we may confine attention to modes associated with real exponential time-factors, and limiting conditions of stability are in fact obtained when all time variations are made zero.

SCOPE OF THE PRESENT PAPER

5. A point of novelty in the present paper is the standpoint adopted in regard to the shape of the cell wall (§ 1). Rayleigh's treatment is restricted to rectangular cells, on the ground that 'by Fourier's theorem the motion in its earlier stages may be analysed into components, each of which corresponds to rectangular cells whose sides are fixed axes arbitrarily chosen'. On the other hand he asserts that 'on a more general view...the disturbance which develops most rapidly may be assimilated to...the free vibration of an infinite stretched membrane vibrating with given frequency', and he makes some approach to a theory of the hexagonal cells which appear from Bénard's experiments to be the final and permanent regime, by treating the hexagon as a *slight* deviation from the circular form He states in regard to cells having the forms of equilateral triangles or hexagons. 'I am not aware that an analytical solution has been obtained for these cases.' Jeffreys (1926, 1928) follows him in assuming rectangular cells for the purpose of a mathematical treatment. By contrast, in the analytical part of this paper the shape of the cell is left indefinite, its influence on the motion being represented by a parameter a^2 which can be interpreted on the basis of the 'membrane analogue' noticed by Rayleigh in the sentence cited above. a is proportional to the gravest natural frequency of a uniform membrane having the shape of the cell in plan and vibrating transversely under appropriate edge conditions. In this way the main problem is reduced to that of relating the 'characteristic number' $-\beta\gamma h^4/k\nu$ (§ 20) with a^2, and the separate problem of evaluating a^2 for any specified shape of cell is left for approximate treatment by 'relaxation methods'. A suitable technique has been evolved by D. G. Christopherson, who was led by this enquiry to discover an *exact* functional solution for the hexagonal cell. His work is cited in § 30.

6. Not only in regard to cell shape but also in the variety of the contemplated boundary conditions this investigation has a wider range than Rayleigh's. He for simplicity assumed that horizontal velocities can occur (although vertical motion is prevented) at both the top and bottom surfaces, whereas at a solid boundary such velocities are excluded by the requirement

of 'no slip'; consequently the conditions of his solution are unreal, unless (possibly) in relation to meteorology. Jeffreys dealt also with the cases (1) of two rigid and conducting boundaries and (2) of a bottom boundary rigid and conducting, with a free non-conducting surface at the top. For reasons stated earlier (§ 1 and footnote), case (2) would seem to relate to a merely temporary instability.

In this paper uniform and constant temperature is postulated at each of the horizontal surfaces, but these may either be free or in contact with rigid surfaces preventing slip. Solutions are obtained for a range of values for a^2 (§ 5), and are believed to be correct to five significant figures. The treatment though laborious is not difficult, and in a preliminary exploration it can be replaced by an alternative method (believed to be new) which is justified in § 34 by the proving of a stationary property analogous with that of which Lord Rayleigh made wide application in the theory of vibrations. Concluding sections of the paper deal with the supplementary determination of θ and of the horizontal component velocities u and v.

7. Grateful acknowledgement is made of help received in the preparation of diagrams from the Secretary and staff of the Aeronautical Research Committee.

GENERAL THEORY

The governing equations

8. In rectangular co-ordinates (Ox, Oy lying in the bottom horizontal plane and Oz being directed vertically upwards) the complete equations of motion are

$$\rho\frac{D}{Dt}(u,v,w)=\rho(X,Y,Z)-\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)\mathrm{p}+\nu\rho\nabla^2(u,v,w) \tag{1}$$

in the customary notation (Lamb 1924, § 328), and in this problem we have

$$X=0,\quad Y=0,\quad Z=-g. \tag{2}$$

The equation of continuity is (Lamb 1924, § 7)

$$\frac{D\rho}{Dt}+\rho\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}\right)=0, \tag{3}$$

and the equation of thermal expansion is

$$\rho=\rho_0(1-\alpha\theta) \tag{4}$$

when the fluid (assumed incompressible) has density ρ_0 at the temperature from which θ is measured, α denoting its coefficient of expansion. Finally, the equation of conduction for heat is

$$\frac{D\theta}{Dt} = k\nabla^2\theta, \tag{5}$$

k denoting the diffusivity for temperature.

9. In the initial (steady) state u, v, w vanish severally, so that the first two of (1) require that $\mathbf{p} = p_0$ shall be independent of x and y, and the third reduces to

$$0 = -g\rho - \frac{\partial p_0}{\partial z}, \tag{i}$$

p_0 denoting the initial (steady) pressure at (x, y, z). Equation (3) is then satisfied identically.

The temperature is steady and independent of x and y, so that (5) reduces to

$$0 = k\frac{\partial^2\theta}{\partial z^2}. \tag{ii}$$

Therefore we may express the initial temperature θ_0 in the form

$$\theta_0 = \Theta_0 + \beta z, \tag{6}$$

β being a measure of the 'steady temperature gradient'. Then Θ_0 is the temperature at the bottom horizontal surface ($z = 0$), and at the top surface

$$z = h \text{ (say)}, \quad \theta_0 = \Theta_0 + \beta h = \Theta_1 \text{ (say)}. \tag{7}$$

Again, in the steady state we have from (4)

$$\rho = \rho_0(1 - \alpha\theta_0), \tag{iii}$$

or, if ρ_0 be now defined as the density corresponding with Θ_0,

$$\rho = \rho_0\{1 - \alpha(\theta_0 - \Theta_0)\} = \rho_0(1 - \alpha\beta z), \text{ by (6)}. \tag{iv}$$

Therefore according to (i)

$$\frac{\partial p_0}{\partial z} = -g\rho_0(1 - \alpha\beta z). \tag{8}$$

10. *Now let u, v, w, the velocities in the convective motion, be assumed sufficiently small to justify neglect of their squares and products:* then in (1) we

may replace D/Dt by $\partial/\partial t$. The consequent change of temperature will also be small, so that writing

$$\boldsymbol{\theta} = \theta_0+\theta = \Theta_0+\beta z+\theta \text{ by } (6) \tag{v}$$

for the temperature as modified by convection, we may neglect terms of the second order in u, v, w, θ. Substituting from (v) in (5), on this understanding we deduce that

$$-\beta w = \left[\frac{\partial}{\partial t}-k\nabla^2\right]\theta; \tag{9}$$

also we have from (4), ρ_0 being defined in (iv) of § 9,

$$\rho = \rho_0\{1-\alpha(\boldsymbol{\theta}-\Theta_0)\} = \rho_0\{1-\alpha(\beta z+\theta)\}. \tag{vi}$$

Comparing (vi) with (iv) we see that the increment in ρ which results from the convective motion is a fraction

$$-\frac{\alpha\theta}{1-\alpha\beta z} \tag{vii}$$

of the steady value, and so (on the above convention) may be neglected when it is multiplied by u, v, w or θ. Similarly, writing

$$\mathbf{p} = p_0+p \tag{viii}$$

for the pressure as modified by convection, we deduce from (1), in virtue of (8), the three equations

$$\left.\begin{aligned} \frac{\partial}{\partial t}(u,v,w) &= (0,0,\gamma\theta)-\frac{1}{\rho}\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)p+\nu\nabla^2(u,v,w),\\ \text{in which}\qquad \gamma &= g\alpha \end{aligned}\right\} \tag{10}$$

and θ and p have the meanings given above.

In (10) ρ stands for the density in the steady state as given by (iv) of § 9, so depends, strictly speaking, both on z and Θ_0. But a similar argument shows that its non-constant part may be neglected when (as here) it is associated with p, and thus in (10) without sensible inaccuracy we may treat ρ as constant. In other words, we may treat α as negligible except in association with g: i.e. we may (with Boussinesq and Rayleigh) disregard in this problem the change of density due to temperature, 'except in so far as it modifies the operation of *gravity*'.*

Consistently with this approximate treatment we may replace (3) by

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z} = 0, \tag{11}$$

---

* Rayleigh 1916, p. 436.

since the quantity which is then neglected, namely

$$\frac{1}{\rho}\frac{D\rho}{Dt},$$

vanishes with α when we substitute for ρ from (vi).

11. Eliminating u and v by combining (11) with the first and second of (10), we have

$$\left[\frac{\partial}{\partial t} - \nu\nabla^2\right]\frac{\partial w}{\partial z} = \frac{1}{\rho}\nabla_1^2 p, \tag{ix}$$

∇_1^2 standing for $\partial^2/\partial x^2 + \partial^2/\partial y^2$, and ρ being again treated as constant. Then eliminating p between this equation and the third of (10), we find that

$$\left[\frac{\partial}{\partial t} - \nu\nabla^2\right]\nabla^2 w = \gamma\nabla_1^2\theta. \tag{12}$$

Another relation between w and θ has been given in (9) of § 10 If between (9) and (12) we eliminate θ, there results

$$\left[\left(\frac{\partial}{\partial t} - \nu\nabla^2\right)\left(\frac{\partial}{\partial t} - k\nabla^2\right)\nabla^2 + \beta\gamma\nabla_1^2\right] w = 0. \tag{13}$$

Eliminating w we obtain a like equation in θ.

12. If we can solve for w, then we can deduce θ from (12) combined with boundary conditions which have still to be discussed. We can also deduce u and v, since from the first and second of (10) we may eliminate p to obtain

$$\left[\frac{\partial}{\partial t} - \nu\nabla^2\right]\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) = 0, \tag{x}$$

—an equation which can be solved by writing

$$\left.\begin{aligned} &u = -\frac{\partial\phi}{\partial x} - \frac{\partial\psi}{\partial y}, \quad v = -\frac{\partial\phi}{\partial y} + \frac{\partial\psi}{\partial x}, \\ &\text{where } \phi \text{ is unrestricted and} \\ &\left[\frac{\partial}{\partial t} - \nu\nabla^2\right]\nabla_1^2\psi = 0. \end{aligned}\right\} \tag{14}$$

Then equation (11) requires that

$$\frac{\partial w}{\partial z} = -\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = \nabla_1^2\phi, \tag{15}$$

and (14) and (15), in combination with appropriate boundary conditions, serve to determine ϕ and ψ.

Separation of the variables. The membrane analogy

13. We are concerned (§ 1) with cells which are bounded by a vertical cylinder and by two horizontal surfaces, so an obvious first step to the solution of (13) is the separation of the variables x and y from z by the assumption that w has the form $f(x,y).F(z).\Phi(t)$, where $\nabla_1^2 f \propto f$ and $\Phi(t)$ is an exponential function. Accordingly we now impose the condition

$$h^2\nabla_1^2 w + a^2 w = 0, \tag{16}$$

in which h denotes the depth of a cell and a^2 is a 'characteristic number' not yet determined.

As remarked by Rayleigh (cf. § 5), an equation of the form of (16) governs the displacement (w) of a uniform membrane having the cell wall as boundary and vibrating freely in a 'normal mode'. If (§ 1) the cell wall is a material boundary on which the fluid cannot slip, then w must be kept zero at the boundary: if it is a surface of symmetry (as assumed by Rayleigh), then the normal gradient $\partial w/\partial n$ must vanish in virtue of that symmetry. In either event a like h must be real, since we have by Green's theorem, according to (16),

$$\begin{aligned} a^2\iint w^2\,dx\,dy &= -h^2\iint w\nabla_1^2 w\,dx\,dy \\ &= h^2\left[\iint\left\{\left(\frac{\partial w}{\partial x}\right)^2+\left(\frac{\partial w}{\partial y}\right)^2\right\}dx\,dy-\oint w\frac{\partial w}{\partial n}\,ds\right]: \end{aligned} \tag{17}$$

the line integral is zero and both of the surface integrals are necessarily positive, so a^2 is a positive quantity.

The boundary conditions

14. At every boundary physical conditions are imposed both on the temperature (θ) and on the fluid velocities (u, v, w); but every condition can be expressed in terms of either quantity, since θ is related with w both by (9) and (12).

Writing n for the normal to the cylindrical boundary drawn outwards, we have at every point of that boundary:

If this is a surface of symmetry (as assumed by Rayleigh)

$$\frac{\partial w}{\partial n} = 0 \quad\text{and}\quad \frac{\partial \theta}{\partial n} = 0; \tag{18}$$

if this is a non-conducting rigid surface on which the fluid cannot slip (e.g. in experiment: cf. § 1 and footnote)

$$w = 0 \quad \text{and} \quad \frac{\partial\theta}{\partial n} = 0 \quad \text{as before.} \tag{19}$$

The condition imposed on w serves to define the parameter a^2 of § 13. The condition $\partial\theta/\partial n = 0$, imposed at every point of the cell wall (i.e. for all values of z), requires according to (9) and (12) that

$$\beta\gamma\frac{\partial w}{\partial n} = k\gamma\frac{\partial}{\partial n}\nabla_1^2\theta = k\frac{\partial}{\partial n}\left[\frac{\partial}{\partial t} - \nu\nabla^2\right]\nabla^2 w. \tag{20}$$

If w satisfies equation (16) this relation is satisfied identically when $\partial w/\partial n = 0$ at every point of the cell wall, consequently the second condition (18) is equivalent to the first. When the cell wall is a non-conducting rigid boundary at which $w = 0$, the second of (19) can be replaced by (20).

15. The top and bottom surfaces may be either free or constrained by rigid boundaries, but in either event are assumed in this paper (cf. § 2) to be maintained at constant temperature. At a free surface w and $\partial u/\partial z$, $\partial v/\partial z$ will vanish, therefore $\partial^2 w/\partial z^2$ in virtue of (11), at a rigid surface on which the fluid cannot slip, w must vanish with u and v, i.e. with $\partial w/\partial z$; on a surface kept at constant temperature we have $\boldsymbol{\Theta} = \theta_0$ always, therefore $\theta = 0$. Accordingly at these surfaces we may have either of the following sets of boundary conditions:

If the surface is free

$$w = D^2 w = 0 \quad \text{and} \quad \theta = 0, \tag{21}$$

if the surface is defined by a rigid boundary

$$w = Dw = 0 \quad \text{and} \quad \theta = 0, \tag{22}$$

D denoting $\partial/\partial\zeta$, where $\zeta = z/h$.

As in § 14 we can express the condition $\theta = 0$ as a third condition to be satisfied by w. For since $\nabla_1^2\theta$ will vanish with θ, according to (12) we have

$$\left[\frac{\partial}{\partial t} - \nu\nabla^2\right]\nabla^2 w = 0 \tag{i}$$

at each surface, whether it be free or constrained by a rigid boundary. Also from (9), when $w = \theta = 0$, it follows that $\nabla^2\theta$ vanishes, therefore $\nabla^2\nabla_1^2\theta$; and so according to (12), at either type of boundary

$$\left[\frac{\partial}{\partial t} - \nu\nabla^2\right]\nabla^4 w = 0. \tag{ii}$$

This, however, is not a new condition, being a consequence of (i) and of the first of (21) or (22) according to the governing equation (13).

Combined with the first and second of (21), equations (i) and (ii) show that at a *free surface*

$$w = D^2w = D^4w = D^6w = \dots \text{ etc.} = 0. \tag{23}$$

At a rigid boundary, (i) replacing the third of (22), we have

$$w = Dw = \left[\frac{\partial}{\partial t} - \nu\nabla^2\right]\nabla^2 w = 0. \tag{24}$$

Either set of conditions, in combination with the governing equation (13), will serve to determine w, and until solutions in w have been obtained we need not again consider θ.

Nature of the exponential time factor $\Phi(t)$. The possibility of oscillatory convective motion

16. When $\Phi(t) \propto e^{\sigma t}$ as assumed in § 13, according to (9) and (12)

$$[\sigma - k\nabla^2]\,\theta = -\beta w, \quad [\sigma - \nu\nabla^2]\,\nabla^2 w = \gamma\nabla_1^2\theta. \tag{25}$$

Eliminating θ between these relations in the manner of § 11, we have

$$[\sigma^2\nabla^2 - (k+\nu)\,\sigma\nabla^4 + k\nu\nabla^6 + \beta\gamma\nabla_1^2]\,w = 0, \tag{26}$$

which is the form assumed by (13) when $\partial/\partial t \equiv \sigma$. We now examine the possibility that σ may take complex or imaginary values.

A complex or imaginary value of σ will be associated with a complex form of w in (26), and its conjugate σ', in association with w' the conjugate of w, will also satisfy that equation together with the boundary conditions (18) or (19), (23) or (24). Moreover according to (25) θ will be complex with w, and its conjugate θ', in association with σ' and w', will also satisfy equations (25) together with the boundary conditions imposed on θ: that is to say, corresponding with (25) we shall have the relations

$$[\sigma' - k\nabla^2]\,\theta' = -\beta w', \quad [\sigma' - \nu\nabla^2]\,\nabla^2 w' = \gamma\nabla_1^2\theta'. \tag{27}$$

Then from the first of (25) and second of (27), by cross-multiplication, there will result

$$\beta w[\sigma' - \nu\nabla^2]\,\nabla^2 w' + \gamma\nabla_1^2\theta'[\sigma - k\nabla^2]\,\theta = 0. \tag{28}$$

17. Now let the left-hand side of this equation be integrated throughout the volume of one cell. Having regard to the boundary conditions (18)–(22),

which are satisfied by w, w' and by θ, θ', and taking account of (16), with the aid of Green's transformation we find that

$$\iiint\left(w\nabla^2 w' + \frac{\partial w}{\partial x}\frac{\partial w'}{\partial x} + \frac{\partial w}{\partial y}\frac{\partial w'}{\partial y} + \frac{\partial w}{\partial z}\frac{\partial w'}{\partial z}\right)d(\text{vol.}) = \iint w\frac{\partial w'}{\partial \nu}\,dS = 0,$$

$$\iiint(w\nabla^4 w' - \nabla^2 w\nabla^2 w')\,d(\text{vol.}) = \iint\left(w\frac{\partial}{\partial \nu}\nabla^2 w' - \frac{\partial w}{\partial \nu}\nabla^2 w'\right)dS = 0,$$

$$\iiint\left(\theta\nabla_1^2\theta' + \frac{\partial \theta}{\partial x}\frac{\partial \theta'}{\partial x} + \frac{\partial \theta}{\partial y}\frac{\partial \theta'}{\partial y}\right)d(\text{vol.}) = \iint \theta\frac{\partial \theta'}{\partial n}\,ds\,dz = 0,$$

$$\begin{aligned}\iiint\left(\frac{\partial^2\theta}{\partial z^2}\nabla_1^2\theta' - \frac{\partial^2\theta}{\partial x\partial z}\frac{\partial^2\theta'}{\partial x\partial z} - \frac{\partial^2\theta}{\partial y\partial z}\frac{\partial^2\theta'}{\partial y\partial z}\right)&d(\text{vol.})\\ = \iiint\left\{\frac{\partial}{\partial z}\left(\frac{\partial\theta}{\partial z}\nabla_1^2\theta'\right) - \frac{\partial}{\partial x}\left(\frac{\partial\theta}{\partial z}\frac{\partial^2\theta'}{\partial x\partial z}\right) - \frac{\partial}{\partial y}\left(\frac{\partial\theta}{\partial z}\frac{\partial^2\theta'}{\partial y\partial z}\right)\right\}&d(\text{vol.})\\ = \iint\left[\frac{\partial\theta}{\partial z}\nabla_1^2\theta'\right]_0^1 dx\,dy - \iint\frac{\partial\theta}{\partial z}\frac{\partial}{\partial n}\left(\frac{\partial\theta'}{\partial z}\right)ds\,dz &= 0,\end{aligned}$$

where as in § 13 s denotes the boundary of a horizontal section of the cell wall and n the outward normal to that boundary. (Thus $ds\,dz = dS_c$, an element of the cell wall, and $dx\,dy = dS_h$, the element of a horizontal surface.) dS denotes, indifferently, either dS_c or dS_h, and ν denotes the outward-drawn normal to dS, in a surface integration extending to the whole cell boundary. Suffixes 0 and 1 relate respectively to the lower and upper boundary.

In virtue of those relations, the result of the integration performed on (28) may be written as

$$\beta\sigma' I_1 + \beta\nu I_2 + \gamma\sigma I_3 + k\gamma I_4 = 0, \tag{29}$$

where

$$\left.\begin{aligned}&I_1 \text{ stands for the integral } \iiint\left(\frac{\partial w}{\partial x}\frac{\partial w'}{\partial x} + \frac{\partial w}{\partial y}\frac{\partial w'}{\partial y} + \frac{\partial w}{\partial z}\frac{\partial w'}{\partial z}\right)d(\text{vol.}),\\ &I_2 \text{ stands for the integral } \iiint\nabla^2 w\nabla^2 w'\,d(\text{vol.}),\\ &I_3 \text{ stands for the integral } \iiint\left(\frac{\partial\theta}{\partial x}\frac{\partial\theta'}{\partial x} + \frac{\partial\theta}{\partial y}\frac{\partial\theta'}{\partial y}\right)d(\text{vol.}),\\ &I_4 \text{ stands for the integral } \iiint\left(\nabla_1^2\theta\nabla_1^2\theta' + \frac{\partial^2\theta}{\partial x\partial z}\frac{\partial^2\theta'}{\partial x\partial z} + \frac{\partial^2\theta}{\partial y\partial z}\frac{\partial^2\theta'}{\partial y\partial z}\right)d(\text{vol.}).\end{aligned}\right\} \tag{30}$$

18. If we had started (§ 16) by cross-multiplication between the second of (25) and first of (27), the relation corresponding with (29) would have been

$$\beta\sigma I_1+\beta\nu I_2+\gamma\sigma' I_3+k\gamma I_4=0, \tag{31}$$

as is evident from the symmetry of the expressions (30). Now writing

$$\left.\begin{aligned}\sigma&=p+\iota q\\ \sigma'&=p-\iota q\end{aligned}\right\} \tag{32}$$

in accordance with § 16, we have by addition of (29) and (31)

$$\left.\begin{aligned}p(\beta I_1+\gamma I_3)+\beta\nu I_2+k\gamma I_4&=0,\\ \text{and by subtraction}\qquad q(\gamma I_3-\beta I_1)&=0,\end{aligned}\right\} \tag{33}$$

in which the I's are all positive according to (30), since every integrand consists of one or more products of conjugate complex quantities.

Since ν and k are positive by definition, the first of (33) shows that p will be negative (i.e. the time factor $\Phi(t)$ will be one of decay) unless $\beta\gamma$ is negative; and if $\beta\gamma$ is negative q must be zero according to the second of (33), i.e. $\Phi(t)$ must be exponential with real index pt. Consequently, while oscillatory motions are not excluded by this investigation, they are permitted only in circumstances making for stability, i.e. in which they decay. Whenever *maintained* convective motion is possible, σ in § 16 must be real: therefore limiting conditions of stability are in fact obtained (§ 4) when all time variations are made zero.

Critical conditions. The simplified equations

19. Accordingly we now proceed on the assumption that $\sigma\equiv\partial/\partial t$ is zero, so that (13) and (26) are replaced by

$$[k\nu\nabla^6+\beta\gamma\nabla_1^2]\,w=0,$$

or

$$[k\nu(D^2-a^2)^3-\beta\gamma h^4a^2]\,w=0 \tag{i}$$

according to (16), D denoting $\partial/\partial\zeta$ as in § 15. Similarly equation (14) is now replaced by

$$\nabla^2\nabla_1^2\psi=0, \tag{14A}$$

and (24) by

$$w=Dw=(D^2-a^2)^2w=0. \tag{24A}$$

For convenience we shall write (i) in the form

$$[(D^2-a^2)^3+\lambda^3a^6]\,w=0, \tag{34}$$

where

$$\lambda^3=-\frac{\beta\gamma h^4}{k\nu a^4}. \tag{35}$$

Our immediate problem is to relate λ with a and thereby to find the lowest value of the 'characteristic number' $-\beta\gamma h^4/k\nu$ for which steady slow motion can occur.

20. The symbol λ was used by Jeffreys to denote this characteristic number: Rayleigh, using γ and ν in their present connotation, writes β' for $-\beta$, ζ for h, κ for k. Following Rayleigh we may obtain a physical interpretation by writing $(\Theta_1-\Theta_0)/h$ for β, $g(\rho_1-\rho_0)/\rho_0(\Theta_0-\Theta_1)$ for γ, where Θ_1, ρ_1 and Θ_0, ρ_0, as in § 9, stand for the temperatures and densities in the initial (steady) regime at the top and bottom surfaces respectively. Then the characteristic number is given by

$$\lambda^2 a^4 = \frac{-\beta\gamma h^4}{k\nu} = \frac{gh^3}{k\nu}\frac{\rho_1-\rho_0}{\rho_0}, \tag{36}$$

in which $(\rho_1-\rho_0)/\rho_0$ is the fractional excess of density in the fluid at the top as compared with the fluid at the bottom (heated) surface. When α and therefore γ is negative, to maintain convective motion heat must be supplied at the top surface,* but the characteristic number will still be given by (36).

21. We have shown that λ^2 must be positive, since it was proved in § 18 that p must be negative unless $\beta\gamma$ is negative, and the other factors of λ^2 (namely, k, ν and h^4) are positive. Using (34) we can proceed further: for on multiplying that equation by $(D^2-a^2)^2 w$ and then integrating with respect to ζ in the range $0 \leqslant \zeta \leqslant 1$ (i.e. between the lower and upper horizontal surfaces) we have

$$\lambda^2 a^6 \int_0^1 w.(D^2-a^2)^2 w\,d\zeta = -\int_0^1 (D^2-a^2)^2 w\ \ (D^2-a^2)^3 w\,d\zeta, \tag{i}$$

D standing for $\partial/\partial\zeta$ as before, also

$$\int_0^1 [w.(D^2-a^2)^2 w - \{(D^2-a^2)w\}^2]\,d\zeta = \Big[w.D(D^2-a^2)w - Dw.(D^2-a^2)w\Big]_0^1, \tag{ii}$$

$$\int_0^1 [(D^2-a^2)^2 w.(D^2-a^2)^3 w + \{D(D^2-a^2)^2 w\}^2 + a^2\{(D^2-a^2)^2 w\}^2]\,d\zeta = \Big[(D^2-a^2)^2 w.D(D^2-a^2)^2 w\Big]_0^1, \tag{iii}$$

and the quantities on the right of (ii) and (iii) vanish at both limits, in virtue either of the boundary conditions (23) or (24)A. Consequently (i) is equivalent to

---

* Cf. footnote to § 1.

$$\lambda^3 a^6 \int_0^1 \{(D^2-a^2)\,w\}^2 d\zeta = \int_0^1 [\{D(D^2-a^2)^2 w\}^2 + a^2\{(D^2-a^2)^2 w\}^2]\, d\zeta, \quad (37)$$

and both integrals in this equation are positive.

EXACT SOLUTIONS OF THE SIMPLIFIED EQUATIONS

22. Having shown that λ is positive, we may solve (34) symbolically in the form

$$D^2 \equiv a^2\{1-\lambda \times (\omega_1, \omega_2, \omega_3)\}, \quad (38)$$

where $\omega_1, \omega_2, \omega_3 = 1, \frac{-1 \pm i\sqrt{3}}{2}$ are the three cube roots of unity, so that

$$\left.\begin{aligned} &\omega_1, \omega_2, \omega_3 = 1, \quad \frac{-1 \pm i\sqrt{3}}{2} \\ &\omega_2^2 = \omega_3, \quad \omega_3^2 = \omega_2. \end{aligned}\right\} \quad (39)$$

The most general solution can then be written in the form

$$\begin{aligned} w = {} & A_1 \cosh 2\mu_1\zeta + A_2 \cosh 2\mu_2\zeta + A_3 \cosh 2\mu_3\zeta \\ & + B_1 \sinh 2\mu_1\zeta + B_2 \sinh 2\mu_2\zeta + B_3 \sinh 2\mu_3\zeta, \quad (40) \end{aligned}$$

where A_1, A_2, A_3, B_1, B_2, B_3 are arbitrary and $4\mu_1^2$, $4\mu_2^2$, $4\mu_3^2$ are the three values of D^2 as given by (38). There are six arbitrary constants, therefore three conditions can be satisfied at each of the horizontal boundaries; but except for particular values of λ the solution will be nugatory in that all six constants are zero.

For convenience we shall now change the origin of ζ so that the lower and upper boundaries are defined by $\zeta = -\frac{1}{2}$, $\zeta = +\frac{1}{2}$ and the range of ζ is $-1 \leqslant 2\zeta \leqslant 1$. Then in (40) we can separate ‘even’ and ‘odd solutions’ in the two symmetrical cases, viz.

(a) that of *two free surfaces* (Rayleigh's case: cf. § 6), in which the boundary conditions (23) must be satisfied when $2\zeta = \pm 1$,

(b) that of *two rigid boundaries*, in which the conditions (24)A must be satisfied when $2\zeta = \pm 1$.

In our third case (c) *one surface is free, the other constrained by a rigid boundary*, so (23) must be satisfied at one and (24)A at the other of the limits $2\zeta = \pm 1$. Here, by reason of the asymmetry, all six terms will appear in (40); but we shall find that in fact solutions for case (c) can be based on those for case (b).

Case (a). Two free surfaces.

23. The solution for this case is simple. The 'even' solutions are derived from

$$w = A_1 \cosh 2\mu_1 \zeta + A_2 \cosh 2\mu_2 \zeta + A_3 \cosh 2\mu_3 \zeta, \tag{41}$$

where by (23), since $\mu_1 \neq \mu_2 \neq \mu_3$,

$$A_1 \cosh \mu_1 = A_2 \cosh \mu_2 = A_3 \cosh \mu_3 = 0.$$

Hence we have solutions typified by

$$w = A \cosh in\pi\zeta = A \cos n\pi\zeta, \tag{42}$$

A being arbitrary and n having any (non-zero) integral value; and on substituting from (42) in (34) we deduce that

$$\lambda^3 = \left(1 + \frac{n^2\pi^2}{a^2}\right)^3. \tag{43}$$

The 'odd' solutions are found in the same way to be typified by

$$w = -iB \sinh 2im\pi\zeta = B \sin 2m\pi\zeta, \tag{44}$$

B being arbitrary and m having any (non-zero) integral value; and we have

$$\lambda^3 = \left(1 + \frac{4m^2\pi^2}{a^2}\right)^3. \tag{45}$$

Case (b). Two rigid boundaries.

24. Here again even and odd solutions can be separated and the even solutions are given by

$$w = A_1 \cosh 2\mu_1 \zeta + A_2 \cosh 2\mu_2 \zeta + A_3 \cosh 2\mu_3 \zeta, \tag{41) \textit{bis}$$

but now, since (24) A must be satisfied when $2\zeta = \pm 1$, we have

$$A_1 \cosh \mu_1 + A_2 \cosh \mu_2 + A_3 \cosh \mu_3 = 0,$$

$$\mu_1 A_1 \sinh \mu_1 + \mu_2 A_2 \sinh \mu_2 + \mu_3 A_3 \sinh \mu_3 = 0,$$

$$(4\mu_1^2 - a^2)^2 A_1 \cosh \mu_1 + (4\mu_2^2 - a^2)^2 A_2 \cosh \mu_2 + (4\mu_3^2 - a^2)^2 A_3 \cosh \mu_3 = 0,$$

showing that A_1, A_2, A_3 will vanish severally unless

$$\begin{vmatrix} \cosh \mu_1, & \cosh \mu_2, & \cosh \mu_3, \\ \mu_1 \sinh \mu_1, & \mu_2 \sinh \mu_2, & \mu_3 \sinh \mu_3, \\ \omega_1^2 \cosh \mu_1, & \omega_2^2 \cosh \mu_2, & \omega_3^2 \cosh \mu_3 \end{vmatrix} = 0. \tag{46}$$

(The third line of the determinant has been simplified after substitution for $\mu_1^2, \mu_2^2, \mu_3^2$ in accordance with their definitions in § 22.)

$\omega_1^2, \omega_2^2, \omega_3^2$ having the values (39), since μ_1, μ_2, μ_3 are non-zero an equivalent condition is

$$\begin{vmatrix} 1, & 1, & 1, \\ \mu_1 \tanh \mu_1, & \mu_2 \tanh \mu_2, & \mu_3 \tanh \mu_3, \\ 0, & \sqrt{3}+i, & \sqrt{3}-i \end{vmatrix} = 0,$$

i.e.
$$2i\mu_1 \tanh \mu_1 + (\sqrt{3}-i)\mu_2 \tanh \mu_2 - (\sqrt{3}+i)\mu_3 \tanh \mu_3 = 0. \qquad (47)$$

25. Now writing

$$\left.\begin{aligned} & 2\mu_2 = (A-iB)a, \quad 2\mu_3 = (A+iB)a, \\ & \text{we have from the definition of } \mu_2, \mu_3 \text{ in § 22} \\ & A^2 - B^2 = 1 + \tfrac{1}{2}\lambda, \quad 2AB = \frac{\sqrt{3}}{2}\lambda, \\ & \text{and hence by elimination of } \lambda \\ & A^2 - B^2 - \frac{2}{\sqrt{3}}AB = 1. \end{aligned}\right\} \qquad (48)$$

Moreover we have

$$2\mu_1 = ia\sqrt{(\lambda-1)} = ia\sqrt{\{(B+\sqrt{3}A)^2 - 4A^2\}}, \qquad (49)$$

so (47) can be written in the form

$$-\sqrt{(\lambda-1)}\tan \tfrac{1}{2}a\sqrt{(\lambda-1)} = \frac{(A+\sqrt{3}B)\sinh aA + (\sqrt{3}A - B)\sin aB}{\cosh aA + \cos aB}. \qquad (50)$$

In exactly the same way, starting with an odd solution in the form

$$w = B_1 \sinh 2\mu_1\zeta + B_2 \sinh 2\mu_2\zeta + B_3 \sinh 2\mu_3\zeta, \qquad (51)$$

we obtain a condition similar to (47) but with hyperbolic cotangents replacing the hyperbolic tangents. Making the substitutions (48) we can transform this to

$$\sqrt{(\lambda-1)}\cot \tfrac{1}{2}a\sqrt{(\lambda-1)} = \frac{(A+\sqrt{3}B)\sinh aA - (\sqrt{3}A - B)\sin aB}{\cosh aA - \cos aB}, \qquad (52)$$

which is similar to (50).

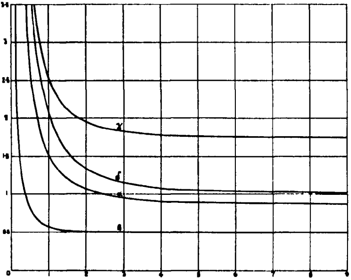
26. Solutions of (50) and of (52) can be obtained by plotting curves to represent the quantities on their left- and right-hand sides. From (48) we have

$$(A+\sqrt{3}B)^2 = 4A^2 - 3,$$

whence (since λ and therefore AB must be positive)

$$\left.\begin{aligned} &\sqrt{3}B = \sqrt{(4A^2-3)} - A \quad \text{and} \quad A^2 > 1. \\ &\text{Then we have} \\ &A + \sqrt{3}B = \sqrt{(4A^2-3)}, \quad \sqrt{3}A - B = \frac{1}{\sqrt{3}}\{4A - \sqrt{(4A^2-3)}\}, \\ &\text{also} \quad \lambda - 1 = \frac{4AB}{\sqrt{3}} - 1 = \frac{4A}{3}\{\sqrt{(4A^2-3)} - A\} - 1, \end{aligned}\right\} \quad (53)$$

by (48). Hence, giving A a series of values in excess of unity, we can calculate, once for all, corresponding values of the coefficients B, $(A+\sqrt{3}B)$, $(\sqrt{3}A - B)$ and $\sqrt{(\lambda-1)}$. The sign of A is immaterial provided that AB is kept positive.



$\sqrt{(\lambda-1)} = 2x/a$

FIGURE 1

$$\text{If now} \quad \left.\begin{aligned} &2x \text{ stands for the quantity } a\sqrt{(\lambda-1)}, \\ &\alpha \text{ stands for the ratio } A/\sqrt{(\lambda-1)}, \\ &\beta \text{ stands for the ratio } B/\sqrt{(\lambda-1)}, \\ &\gamma \text{ stands for the ratio } (A+\sqrt{3}B/\sqrt{(\lambda-1)}, \\ &\delta \text{ stands for the ratio } (\sqrt{3}A - B)/\sqrt{(\lambda-1)}, \end{aligned}\right\} \quad (54)$$

so that x, α, β, γ, δ now have new significance, these results may be recorded in the form of curves (figure 1) which present α, β, γ, δ as functions of

$\sqrt{(\lambda-1)} = 2x/a$. Also equations (50) and (52) can be written in the equivalent forms

$$-\tan x = \frac{\gamma \sinh 2\alpha x + \delta \sin 2\beta x}{\cosh 2\alpha x + \cos 2\beta x} \tag{50 A}$$

and

$$\cot x = \frac{\gamma \sinh 2\alpha x - \delta \sin 2\beta x}{\cosh 2\alpha x - \cos 2\beta x}, \tag{52 A}$$

so can be solved with the aid of curves of $-\tan x$ and $\cot x$, drawn once for all on a base of x. Except when αx is small the expression on the right of either equation will be closely represented by

$$\gamma \tanh 2\alpha x \ (= \gamma, \text{ to six significant figures, when } \alpha x < 4),$$

and when some definite value has been attached to a, either this or the complete expression can be plotted on a base of x. Points of intersection will define roots of (50) A or (52) A.

27. Figure 2 exhibits some solutions obtained in this manner. First, two series of (asymptotic) curves for $-\tan x$ and for $\cot x$ were drawn; then curves of γ corresponding with $a = 2$, 3, 4 and 5. The intersections suggest restricted ranges of x in which roots may be expected to lie, and for those ranges curves of $-\tan x$ and of $\cot x$ were constructed to a much more open scale, also curves representing the complete expression on the right-hand side of (50) A or (52) A.* In this way roots of those equations, regarded as equations in x, were found with an error not greater than 1 in the sixth significant figure.

From the two sets of roots as thus determined, critical values of λ were deduced by means of the relation

$$\lambda = 1 + \frac{4x^2}{a^2}, \tag{55}$$

which is equivalent to the first of (54). Table 1 records the first three of these values as deduced from equation (50) A, table 2 the first three values as deduced from equation (52) A. The last three columns of each table record corresponding values of the 'characteristic number' of § 19, viz.

$$-\frac{\beta \gamma h^4}{k\nu} = \lambda^3 a^4. \tag{56}$$

Only five significant figures are given, since a fractional error of 10^{-6} in x would entail a fractional error of 6×10^{-6} in λ^3.

---

* For ranges in which $2\alpha x > 13$ the complete expressions on the right of (50) A and (52) A were found to agree with γ to six significant figures.

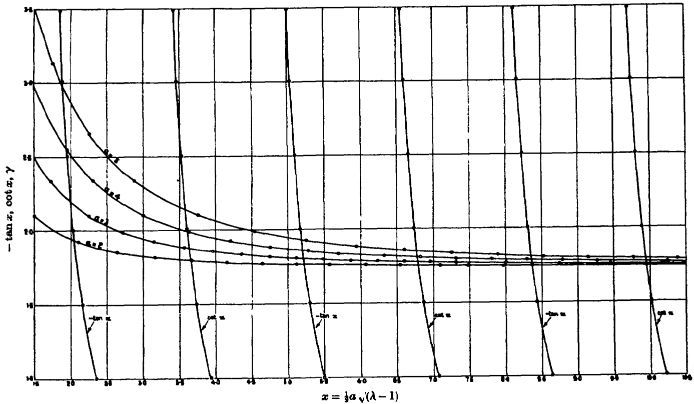

− tan x, cot x, λ
1·0
1·5
2·0
2·5
3·0
3·5
1·5
2·0
2·5
3·0
3·5
4·0
4·5
5·0
5·5
6·0
6·5
7·0
7·5
8·0
8·5
9·0
9·5
9·6
10·0
$a = 5$
$a = 4$
$a = 3$
$a = 2$
−tan x
cot x
−tan x
cot x
−tan x
cot x
$x = \frac{1}{2}a\sqrt{(\lambda - 1)}$

FIGURE 2

TABLE 1. SYMMETRICAL MODES DEDUCED FROM EQUATION (50)

	λ			$-\beta\gamma h^4/k\nu = \lambda^2 a^4$		
	1st mode	2nd mode	3rd mode	1st mode	2nd mode 10 ×	3rd mode 100 ×
$a=2$	5·14367	28·3347	71·1343	2177·4	36398	57591
3	2·76447	13·1045	32·1418	1711·3	18228	26896
4	1·94349	7·77543	18·4953$_5$	1879·3	12034	16197
5	1·57447	5·31047	12·1794	2439·4	93601	11292

TABLE 2. SKEW-SYMMETRICAL MODES DEDUCED FROM EQUATION (52)

	λ			$-\beta\gamma h^4/k\nu = \lambda^2 a^4$		
	1st mode	2nd mode	3rd mode	1st mode	2nd mode 10 ×	3rd mode 100 ×
$a=2$	14·3222	47·2691	99·9269	47006	168990	159650
3	6·85976	21·5295	44·9440	26146	80833	73536
4	4·25234	12·5214	25·6997	19684·5	50257	43453
5	3·04991	8·35258	16·7927	17731	36420	29597

Case (*c*). *One free surface, one rigid boundary.*

28. The 'odd' solutions in case (*b*) are expressions having the form of (51) which satisfy all three of the conditions (24) A when $2\zeta = \pm 1$. In addition, when $\zeta = 0$ they satisfy all three of the conditions (23), because then $\sinh \mu\zeta$, together with its second, fourth and all even derivatives, vanishes without restriction on μ. Accordingly from these odd solutions we can also deduce solutions for case (*c*) of § 22, where one of the horizontal surfaces is 'free', the other constrained by a rigid boundary. But whereas in case (*b*) the two surfaces are defined by $2\zeta = \pm 1$, in case (*c*) the origin of ζ must be in one surface, so the range is halved· this means that any odd solution of case (*b*) which applies to a cell of depth h will apply in case (*c*) to a cell of depth $h' = \frac{1}{2}h$

Because the cell-depth h is related with a by (16), the other (horizontal) cell dimensions will also be different in case (*c*). We shall return to this question when we have dealt with the problem of affixing values to a. What matters for the moment is that no special consideration (on the lines of §§ 24–27) need be given to case (*d*).

Determination of the parameter a (1) *for a rigid cylindrical boundary*

29. The values which in any particular instance can be assumed by a will depend upon the depth and cross-section of the fluid cell, also on the condition which must be satisfied by w at the cylindrical boundary. In experiment the cell wall will be rigid and of specified shape, and the boundary

condition will be $w = 0$; then in the 'membrane analogy' (§ 13) the boundary lies in the cell wall and on it the transverse displacement must vanish. For any shape there will be an indefinite number of possible modes, each associated with some critical value of a: e.g. when the boundary is a circle of radius R the expression

$$w = A_n \cos(n\theta + \alpha_n) J_n(kr) \tag{57}$$

(A_n and α_n being arbitrary constants) will satisfy (16) if

$$k^2 = a^2/h^2, \tag{58}$$

and it will vanish at the boundary ($r = R$) provided that

$$J_n(kR) = 0. \tag{59}$$

The roots of (59) are known, and for each, using (58), we can calculate the corresponding value of a expressed as a fraction of h/R. Table 3 records, for $n = 0, 1, 2, 3$, the four lowest critical values of aR/h appropriate to a rigid circular cell wall. It is based on Table B of § 206 in Chap. IX of Lord Rayleigh's *Theory of Sound*.

TABLE 3. CRITICAL VALUES OF aR/h FOR CIRCULAR CELL

$n =$	0	1	2	3
First	2 404	3 832	5 135	6 379
Second	5 520	7 016	8 417	9 760
Third	8·654	10 173	11·620	13 017
Fourth	11·792	13 323	14·796	16 224

The result obtained in this instance, that a is proportional to the depth h divided by a length (R) which defines the size of the cell in plan, holds also in respect of other shapes and of other boundary conditions. Thus if the cell wall is a rigid square of side L, then a will be proportional to h/L.*

Determination of the parameter a (2) when the cylindrical boundary is a surface of symmetry

30. The boundary condition $\dfrac{dw}{dn} = 0$ (18) *bis*

must be satisfied at the cell wall when this is a surface of symmetry. Only three possibilities† demand attention, namely,

---

* Solutions appropriate to several shapes of boundary will be found in Chap. IX (on the Vibrations of Membranes) of Lord Rayleigh's *Theory of Sound*.

† I.e. of complete symmetry, which requires that n, the number of the sides, shall be expressible in the form $2 + 4/(N-2)$, N being integral.

(1) the boundary may be an equilateral triangle of side L,
(2) the boundary may be rectangular, of sides L_1 and L_2,
(3) the boundary may be a regular hexagon of side L,

and evidently the solutions in case (1) are included in those for case (3).

Case (2) is simple: for (18) will be satisfied (if the origin is taken at the centre of the rectangle) by

$$w = A_{mn} \cos\frac{m\pi x}{L_1} \cos\frac{n\pi y}{L_2} \tag{60}$$

(A_{mn} being constant) when m and n are *any even integers*; and this expression will also satisfy (16) provided that

$$a^2 = h^2\left(\frac{m^2}{L_1^2} + \frac{n^2}{L_2^2}\right). \tag{61}$$

The regular hexagon (case 3) has more interest in relation to our thermal problem, because it appears from Bénard's experiments (cf. § 5) that hexagonal cells characterize the permanent regime in a layer of unlimited extent. The necessity of approximate treatment in the manner of Rayleigh (§ 5) has been removed by the recent discovery of an exact solution (Christopherson 1940). The expression

$$w = \tfrac{1}{3}w_0\left\{\cos\frac{2n\pi}{3L}(\sqrt{3}x+y) + \cos\frac{2n\pi}{3L}(\sqrt{3}x-y) + \cos\frac{4n\pi y}{3L}\right\}, \tag{62}$$

or

$$\left.\begin{aligned} w = \tfrac{1}{3}w_0\Big\{&\cos\frac{2n\pi r}{3L}(\sqrt{3}\cos\theta+\sin\theta) \\ &+\cos\frac{2n\pi r}{3L}(\sqrt{3}\cos\theta-\sin\theta)+\cos\frac{4n\pi r}{3L}\sin\theta\Big\}, \\ \text{where} \qquad & r\cos\theta = x, \quad r\sin\theta = y, \end{aligned}\right\} \tag{63}$$

is symmetrical with respect to the line $\theta = 0$, and it is not altered by an increase of θ to $\theta + \frac{1}{3}\pi$, therefore it is also symmetrical with respect to the lines $\theta = \frac{1}{3}\pi$, $\theta = \frac{2}{3}\pi$. Evidently w_0 is the value of w when $x = y = 0$, i.e. at the centre of the hexagon.

According to (62)

$$\begin{aligned} \frac{dw}{dn} &= \mp\frac{2n\pi}{3\sqrt{3}L}w_0\left\{\sin\frac{2n\pi}{3L}(\sqrt{3}x+y) + \sin\frac{2n\pi}{3L}(\sqrt{3}x-y)\right\} \\ &= -\frac{4n\pi}{3\sqrt{3}L}w_0 \sin n\pi \cos\frac{2n\pi}{3L}y = 0 \text{ if } n \text{ is integral,} \end{aligned}$$

when $x = \pm\frac{\sqrt{3}}{2}L$. So when w is given by (62) then $dw/dn = 0$ (i.e. *the boundary condition is satisfied*) on the regular hexagon whose sides are defined by

$$\left(x \pm \frac{\sqrt{3}}{2}L\right)\left(\frac{x}{\sqrt{3}} + y \pm L\right)\left(\frac{x}{\sqrt{3}} - y + L\right) = 0 \tag{64}$$

and accordingly have length L. Moreover we have from (62)

$$\nabla_1^2 w = -\left(\frac{4n\pi}{3L}\right)^2 w,$$

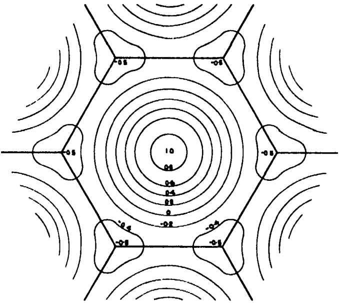


FIGURE 3

and this equation is equivalent to (16) when

$$\frac{aL}{h} = \frac{4n\pi}{3}, \tag{65}$$

n being integral. Accordingly for hexagonal surfaces of symmetry (65) gives the critical values of a.

Figure 3 shows the nature of this important solution, by curves representing contours of the ratio w/w_0.

The criterion of stability

31. In § 19 our problem was reduced to that of finding the least value of

$$-\beta\gamma h^4/k\nu = \lambda^3 a^4 \qquad (56)\ bis$$

for which, under particular conditions imposed at the horizontal surfaces, non-zero solutions of (34) exist, i.e. *steady* solutions of the governing equation (13). The criterion of thermal stability will be that $-\beta\gamma h^4/k\nu$ must not exceed this value.

When the cylindrical boundary is of specified shape and rigid (§ 29), a series of possible values can be attached to a, and hence a series of possible values of $\lambda^3 a^4$ can be determined with the aid of (43) and (45) in case (a) or of tables 1 and 2 in case (b). *The smallest value of $\lambda^3 a^4$ so found is to be inserted in the criterion of stability.*

Suppose, on the other hand, that the fluid has specified depth h but indefinite horizontal extension. Then evidently it has freedom to form a symmetrical cell pattern in which the individual cells are either triangles, rectangles or hexagons, and the size of the cells in plan is not restricted: consequently *all* values of a^2 are admissible, and (according to our approximate theory) the cell pattern which forms will be one which entails the smallest possible value of the 'characteristic number'.

32. In case (a), according to (43) and (45), the least value of $\lambda^3 a^4$ is obtained when $n = 1$ in (43) and when

$$4a^3\lambda^4 + a^4\frac{\partial\lambda^3}{\partial a} = 4\left(a+\frac{\pi^2}{a}\right)^2\left(a-\frac{\pi^2}{2a}\right) = 0,$$

i.e. when $a^2 = \pi^2/2$, so that $\lambda^3 = 27$.

Accordingly the criterion of stability in this case is*

$$-\beta\gamma h^4/k\nu = \lambda^3 a^4 \not> \frac{27\pi^4}{4} = 657\cdot5. \qquad (66)$$

In case (b), inspection of tables 1 and 2 (§ 27) shows that the smallest value of $\lambda^3 a^4$ corresponds (in column 4 of table 1) with a value of a in the neighbourhood of 3. A more detailed exploration led to results which are exhibited in figure 4: the least value of $\lambda^3 a^4$ is 1707·8, and it corresponds with a value of a in the neighbourhood of 3·13. Accordingly the criterion of stability in this case is

$$-\beta\gamma h^4/k\nu = \lambda^3 a^4 \not> 1707\cdot8, \qquad (67)$$

and if this limit is exceeded a cell pattern will form for which $a \approx 3\cdot13$.†

---

* Jeffreys (1926, § 4) and Low (1929, p. 181) both give the numerical value as 651.

† Jeffreys (1928, § 4) gives the figure 1709·5, corresponding with $a = 3\cdot17$. Low (1929, p. 188) obtained 1767, 1756, 1704·4 as successive approximations.

33. Case (*c*) has been noticed in § 28, where it was shown that any odd solution of case (*b*), applicable to a cell of depth h, also provides a solution of case (*c*), applicable to a cell of depth $h' = \frac{1}{2}h$. This means that the last three columns of table 2 record values of

$$-\frac{\beta\gamma h^4}{k} = -16\frac{\beta\gamma h'^4}{k\nu},$$

—that is, 16 times the characteristic number appropriate to case (*c*); also that the values of aR/h which are given in table 3, and the value given in (65) for aL/h, must now be interpreted as values of $aR/2h'$ and of $aL/2h'$ respectively. (Similarly in (61) h^2 must be replaced by $4h'^2$)

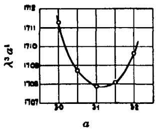


FIGURE 4

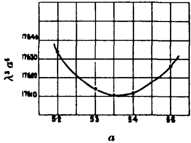


FIGURE 5

Calculations extending the range of table 2 (§ 27) showed that the smallest value assumed by $\lambda^3 a^4$ in relation to 'odd' modes corresponds with a value of a in the neighbourhood of 5, and a more detailed exploration gave results which are exhibited in figure 5 (corresponding with figure 4 for 'even' modes). The least value of $\lambda^3 a^4$ is 17610·5, and in relation to case (*c*), as we have just shown, this quantity stands for 16 times the characteristic number: consequently the criterion of stability in case (*c*) is*

$$-\frac{\beta\gamma h'^4}{k\nu} \ngtr \frac{17610{\cdot}5}{16} = 1100{\cdot}65, \tag{68}$$

and if this value is exceeded a cell-pattern will form for which a (*interpreted as above*) $\approx 5{\cdot}36$.

STATIONARY PROPERTY OF THE CHARACTERISTIC NUMBER

34. We now describe a method, analogous to the use of 'Rayleigh's principle' in vibration theory, which may be used as alternative to the exact treatment given above. It is based upon the integral equation (37) of § 21.

---

* Low (1929, p. 189) obtained 1180 as a first approximation.

Substituting in that equation the correct expression for w, we obtain the true value of λ^3 appropriate to the particular boundary conditions. When on the other hand w is replaced by $w+\delta w$, then the consequent alteration in λ^3 is given as concerns terms of the first order in δw by

$$\left.\begin{aligned}
a^6\delta(\lambda^3) &= \frac{1}{(\text{Den.})}\left[\delta(\text{Num.}) - \left(\frac{\text{Num.}}{\text{Den.}}\right)\delta(\text{Den.})\right]\\
&= \frac{1}{(\text{Den.})}[\delta(\text{Num.}) - \lambda^3 a^6\delta(\text{Den.})], \quad \text{by (37) again,}\\
&\text{where}\\
&(\text{Num.}) \text{ stands for } \int_0^1 \{[D(D^2-a^2)^2 w]^2 + a^2[(D^2-a^2)^2 w]^2\}\,d\zeta,\\
&(\text{Den.}) \text{ stands for } \int_0^1 [(D^2-a^2)\,w]^2\,d\zeta,\\
&\text{so that}\\
\delta(\text{Num.}) &= 2\int_0^1 \{[D(D^2-a^2)^2 w]\,[D(D^2-a^2)^2\,\delta w]\\
&\qquad + a^2[(D^2-a^2)^2 w]\,[(D^2-a^2)^2\,\delta w]\}\,d\zeta,\\
\delta(\text{Den.}) &= 2\int_0^1 [(D^2-a^2)\,w]\,[(D^2-a^2)\,\delta w]\,d\zeta.
\end{aligned}\right\} \qquad (69)$$

Integrating by parts, we may replace the last two of (69) by

$$\left.\begin{aligned}
\delta(\text{Num.}) &= 2\left[\delta Z\frac{\partial Z}{\partial\zeta}\right]_0^1 - 2\int_0^1 \delta Z\,.\,(D^2-a^2)\,Z\,d\zeta,\\
\delta(\text{Den.}) &= 2\left[\frac{\partial w}{\partial\zeta}(D^2-a^2)\,\delta w - w\frac{\partial}{\partial\zeta}(D^2-a^2)\,\delta w\right]_0^1 + 2\int_0^1 w\,\delta Z\,d\zeta,\\
&\text{in which} \quad Z \text{ stands for } (D^2-a^2)^2 w,\\
&\phantom{\text{in which}} \quad \delta Z \text{ stands for } (D^2-a^2)^2\,\delta w,
\end{aligned}\right\} \qquad (70)$$

and then, if the limits of integration are the horizontal boundaries (at which either (23) or (24) A must be satisfied by w and therefore by δw), the first terms on the right of (70) vanish, leaving only the definite integrals. Consequently we may write the first of equations (69) as

$$a^6\delta(\lambda^3) = \frac{-2}{(\text{Den.})}\int_0^1 \delta Z\{(D^2-a^2)\,Z + \lambda^3 a^6 w\}\,d\zeta,$$

and it follows that $\delta(\lambda^3)$ will vanish for all variations δZ, provided that *at every point in the fluid*

$$(D^2-a^2)\,Z + \lambda^3 a^6 w = (D^2-a^2)^3 w + \lambda^3 a^6 w = 0, \qquad (34)\ \textit{bis}$$

as will be the fact when w is an exact solution.

We conclude that not only will λ^3 as calculated from (37) on the basis of an assumed form of w be exact, but it will moreover be stationary in respect of any small variations which are allowed by the boundary conditions, when the assumed form is correct. Consequently,

(*a*) any reasonably close approximation to the correct form for w will, when inserted in (37), lead to a very close estimate of λ^3,

(*b*) in regard to the smallest value of λ^3 an estimate obtained in this way will err, if at all, on the side of *excess*: that is to say, the limit of stability will be placed too high.

These conclusions hold in respect of all of our cases (*a*), (*b*) and (*c*), and without restriction on a, i.e. for all shapes of cylindrical boundary. Either of the conditions (23) or (24) A requires Z (and therefore δZ) to vanish; but between the boundaries δZ is unrestricted, since the solution of

$$(D^2-a^2)^2 w = Z \qquad (70)\ bis$$

can be made to satisfy one or other of the conditions

$$w = D^2 w = 0 \qquad (23)\ bis$$

or

$$w = Dw = 0 \qquad (24)\text{A}\ bis$$

at either boundary.

Application of the stationary property in case (b)

35. To apply the conclusion we assume a form for Z involving one or more arbitrary parameters; we deduce the corresponding form of w according to (70) and (23) or (24) A; and we then adjust the parameters so as to make the quantity

$$\frac{\int_0^1 \{(DZ)^2 + a^2 Z^2\}\, d\zeta}{a^2 \int_0^1 \{(D^2-a^2)\, w\}^2\, d\zeta} = \lambda^3 a^4 \quad \text{according to (37)} \qquad (71)$$

as small as possible. What results will be a close but slightly excessive estimate of the characteristic number.

Suppose, for example, that the boundary conditions are those of case (*b*), that the origin is taken at the centre of the field, and that we are concerned with 'even' solutions (§ 24). Taking as the assumed form

$$\left.\begin{aligned} &Z = A(1 + A_1 \cos \pi\zeta + \cos 2\pi\zeta), \\ &\text{we obtain} \\ &w = A\left[\frac{1}{a^4} + \frac{A_1 \cos \pi\zeta}{(\pi^2+a^2)^2} + \frac{\cos 2\pi\zeta}{(4\pi^2+a^2)^2} + Pa \cosh a\zeta + Qa^2\zeta \sinh a\zeta\right] \end{aligned}\right\} \qquad (72)$$

as the solution of (70) which satisfies (24)A when $2\zeta = \pm 1$, P and Q being constants determined by the terminal conditions

$$\left.\begin{aligned} \frac{1}{a^4} - \frac{1}{(4\pi^2+a^2)^2} + Pa\cosh\frac{a}{2} + \tfrac{1}{2}Qa^2\sinh\frac{a}{2} &= \frac{w}{A} = 0, \\ -\frac{\pi A_1}{(\pi^2+a^2)^2} + Pa^2\sinh\frac{a}{2} + Qa^2\left(\sinh\frac{a}{2} + \frac{a}{2}\cosh\frac{a}{2}\right) &= \frac{Dw}{A} = 0. \end{aligned}\right\} \tag{73}$$

Then $$[D^2-a^2]\,w = -A\left(\frac{1}{a^2} + \frac{A_1\cos\pi\zeta}{\pi^2+a^2} + \frac{\cos 2\pi\zeta}{4\pi^2+a^2} - 2a^3Q\cosh a\zeta\right),$$

and it is easy to show that for the limits $2\zeta = \pm 1$

$$\int (DZ)^2 d\zeta = \pi^2A^2\left(2 + \frac{16}{3\pi}A_1 + \frac{A_1^2}{2}\right),$$

$$a^2\int Z^2 d\zeta = a^2A^2\left(\frac{3}{2} + \frac{16}{3\pi}A_1 + \frac{A_1^2}{2}\right),$$

$$\begin{aligned} a^2\int\{[D^2-a^2]\,w\}^2 d\zeta = a^2A^2\bigg\{&\frac{1}{a^4} - 8Q\sinh\frac{a}{2} + 2Q^2a^5(a+\sinh a) \\ &+ \frac{A_1^2}{2(\pi^2+a^2)^2} + \frac{1}{2(4\pi^2+a^2)^2} + \frac{16A_1(3\pi^2+a^2)}{3\pi a^2(\pi^2+a^2)(4\pi^2+a^2)} \\ &- \frac{8\pi Qa^3}{(\pi^2+a^2)^2}A_1\cosh\frac{a}{2} + \frac{8Qa^4}{(4\pi^2+a^2)^2}\sinh\frac{a}{2}\bigg\}. \end{aligned}$$

But from (73) we have (on elimination of P)

$$\tfrac{1}{2}Qa^2(a+\sinh a) = \frac{\pi A_1}{(\pi^2+a^2)^2}\cosh\frac{a}{2} + \frac{8\pi^2(2\pi^2+a^2)}{a^3(4\pi^2+a^2)^2}\sinh\frac{a}{2},$$

and on making use of this relation we find that

$$\begin{aligned} &a^2\int\{[D^2-a^2]\,w\}^2 d\zeta \\ &\quad= A^2\bigg[\frac{1}{a^2} + \frac{a^2}{2(4\pi^2+a^2)^2} + \frac{A_1^2a^2}{2(\pi^2+a^2)^2} + \frac{16A_1(3\pi^2+a^2)}{3\pi(\pi^2+a^2)(4\pi^2+a^2)} \\ &\qquad - \frac{8a^3}{(a+\sinh a)}\left\{\frac{\pi A_1}{(\pi^2+a^2)^2}\cosh\frac{a}{2} + \frac{8\pi^2(2\pi^2+a^2)}{a^3(4\pi^2+a^2)^2}\sinh\frac{a}{2}\right\}^2\bigg]. \end{aligned}$$

Now, for any assumed value of a, we can deduce from (71) an expression of the form

$$\lambda^3a^4 = \frac{f+gA_1+hA_1^2}{l+mA_1+nA_1^2}, \tag{74}$$

in which f, g, h, l, m, n, are known numerical coefficients, and then the condition that λ^2a^4 shall be stationary in respect of A_1 yields the quadratic equation

$$gl - mf + 2(hl - fn)A_1 + (hm - gn)A_1^2 = 0,$$

from which A_1 can be calculated and inserted in (74).

36. In this way values were obtained as under for comparison with the 'exact' values given in table 1 (§ 27):

TABLE 4

	$-\beta\gamma h^4/k = \lambda^2a^4$	
	'Exact' values (table 1)	Values by approximate method
$a=2$	2177·4	2177·6
3	1711·3	1711·4
4	1879·3	$1879{\cdot}3_5$
5	2439·4	2439·4

The accuracy of the approximate method is remarkable.

COMPLETION OF THE SOLUTION

37. It remains to complete the solution of the simplified equations of small motion (§ 19) when w has been determined in the form $f(x, y) . F(z)$, $\Phi(t)$ in § 13 now having a constant value since all time variations are assumed to vanish (§ 18). $f(x, y)$ is a solution of equation (16) which satisfies the boundary condition (18) or (19).

On this understanding we have from (12) as simplified

$$\gamma\nabla_1^2\theta = -\nu\nabla^4w = \nu\frac{h^2}{a^2}\nabla_1^2\nabla^4w \text{ by (16)},$$

and it follows that

$$\gamma\theta = \nu\frac{h^2}{a^2}\nabla^4w + \chi, \quad \text{where} \quad \nabla_1^2\chi = 0. \tag{i}$$

Then according to (9) as simplified we have

$$\begin{aligned}\beta\gamma w = k\gamma\nabla^2\theta &= \nu k\frac{h^2}{a^2}\nabla^6w + k\frac{\partial^2\chi}{\partial z^2},\\ &= -\beta\gamma\frac{h^2}{a^2}\nabla_1^2w + k\frac{\partial^2\chi}{\partial z^2}, \quad \text{by (13) as simplified,}\\ &= \beta\gamma w + k\frac{\partial^2\chi}{\partial z^2}, \quad \text{by (16) again.}\end{aligned}$$

Consequently χ is linear in respect of z; and because at both horizontal surfaces, by (21) or (22) combined with (12),

$$\theta = 0 \quad \text{and} \quad \nabla^4 w = 0, \quad \text{therefore } \chi = 0 \text{ by (i)},$$

it follows that χ is zero everywhere, so that (i) may be replaced by

$$\gamma\theta = \nu\frac{h^2}{a^2}\nabla^4 w, \tag{ii}$$

$$= -\gamma\frac{h^2}{a^2}\nabla_1^2\theta \quad \text{according to (12).}$$

This means that if equation (16) is satisfied by w, θ will satisfy an equation of identical form. therefore for a particular value of a it will not (in general) be possible to satisfy simultaneously both of the boundary conditions (19). We can on the other hand satisfy both of (18), since according to (ii) and (16) we have

$$\gamma\theta = \frac{\nu}{h^2a^2}(D^2-a^2)^2 w, \tag{75}$$

and when θ is calculated from this expression, on the cylindrical boundary $\partial\theta/\partial n$ will vanish with $\partial w/\partial n$.

We conclude from this investigation that the case of a rigid non-conducting boundary (§§ 1, 14) is not tractable on the basis of a separation of the variables (§ 13). When the cell wall is a surface of symmetry (§ 1), θ comes quite simply from (75).

38. Finally we have to calculate u and v in the case which has been shown to be tractable. When w satisfies (16), we have according to (15)

$$\phi = -\frac{h^2}{a^2}\frac{\partial w}{\partial z} + \phi', \tag{76}$$

ϕ' denoting a plane-harmonic function of x and y. Now at the cylindrical boundary, if this is a surface of symmetry, u and v must satisfy the condition

$$0 = u\cos(x, n) + v\cos(y, n),$$

$$= -\frac{\partial\phi}{\partial n} - \frac{\partial\psi}{\partial s} \quad \text{according to the first and second of (14),}$$

where according to the third of (14), since the motion is steady,

$$\nabla^2\nabla_1^2\psi = 0.$$

We shall satisfy both conditions if ψ = constant and if $\partial\phi/\partial n$ vanishes at every point in the boundary; and when $\partial w/\partial n = 0$ at the boundary, according to (76) this last condition requires that

$$\phi' = \text{const.}$$

Then we have from (14) and (76)

$$u = \frac{h^2}{a^2}\frac{\partial^2 w}{\partial z \partial x}, \quad v = \frac{h^2}{a^2}\frac{\partial^2 w}{\partial y \partial z}. \tag{77}$$

REFERENCES

Bénard, H. 1900 *Rev. Gén. Sci. Pur. Appl.* **12**, 1261, 1309.
— 1901 *Ann. Chim. Phys.* **23**, 62.
Christopherson, D. G. 1940 *Quart. J. Math.* **11**, 63–65.
Jeffreys, H. 1926 *Phil. Mag.* **2**, 833–844.
— 1928 *Proc. Roy. Soc.* A, **118**, 195–208
Lamb, H. 1924 *Hydrodynamics*, 5th ed. Camb Univ. Press.
Low, A. R. 1929 *Proc. Roy Soc.* A, **125**, 180–195.
— 1930 *Proc 3rd Int. Congr. Appl. Mech.* (*Stockholm*).
Rayleigh, Lord 1916 *Phil. Mag.* **32**, 529–546. (Reprinted in *Collected Papers*, **6**, 432–436)

Thermal ionization of strontium

By B. N. Srivastava, D.Sc.
Physics Laboratory, Allahabad University

(*Communicated by M. N. Saha, F.R.S.—Received 4 March* 1940)

The theory of thermal ionization of gases was first given by M. N. Saha in a series of papers (1920*a*, 1920*b*, 1921) and widely applied by him and others to furnish a satisfactory physical theory of stellar spectra. The theory was further developed by Fowler (1923) to include the different excited states of the atom and of the ionized atom, all of which are simultaneously in thermodynamical equilibrium. At the time when the theory was first formulated the ionization potentials and the spectral characteristics of only a few elements were known, but now we possess almost a complete knowledge of the ionization potentials and the spectra of almost all the elements in their different stages of ionization. Hence a detailed comparison of the theory with experimental results can now be satisfactorily undertaken.

The experimental investigation, however, involves the difficult high temperature technique since very high temperatures are necessary to produce a measurable ionization even in the case of easily ionizable gases. Saha, Sur and Majumdar (1927) showed qualitatively from measurements of electrical conductivity that the alkali metals can be ionized by heating their vapours to about 2000° K. The quantitative verification of the theory was first achieved by I. Langmuir and K. H. Kingdon (1925) for caesium vapour by an ingenious method in which caesium atoms striking the high temperature filament of a thermionic valve became ionized. This method was further utilized by T. J. Killian (1926) for rubidium and potassium, by E. Meyer (1930) for potassium and by N. Morgulis (1934) for sodium. The method is, however, complicated by the adsorption of the atoms or ions on the surface of the filament and the thickness of this coating varies with the temperature of the filament, the pressure of the vapour, etc.

Quite a different method has been recently employed by the present author (Srivastava 1940) for investigating the thermal ionization of barium. This method has been found very satisfactory and is being applied in this laboratory to a large number of substances. In the present paper this method has been utilized for studying the thermal ionization of strontium. The demountable vacuum graphite furnace used in these experiments has been described in detail by M. N. Saha and A. N. Tandon (1936) and the internal arrangement employed here has been fully described by the author in the paper on barium. Essentially it consists of a graphite tube heated to about 1600° C by a current of the order of a thousand amperes from a low-tension transformer. This tube has a fine circular hole on one side while on the other side an auxiliary furnace containing solid strontium is fitted gas-tight into it. This auxiliary furnace was thus heated by conduction and the strontium vapour formed there entered the main furnace where it ionizes into Sr^+ ions and electrons. The products of ionization effuse out through the narrow orifice after which they traverse a magnetic field produced by two electromagnets placed on either side. The magnetic field served to show that the negative current almost wholly consists of electrons while the positive particles are much heavier as is expected. The effusion beam is limited by a diaphragm of radius r placed at a distance d from the effusion hole. Behind the diaphragm is a Faraday cylinder connected to a sensitive galvanometer and suitable positive or negative potentials are applied between the Faraday cylinder and the diaphragm in order to collect the particles of the desired charge.

THEORY

Under the conditions it can be shown that the galvanometer current i_g is given by the relation

$$i_g = \frac{epS}{\sqrt{(2m\pi kT)}}\frac{r^2}{r^2+d^2},$$

where S is the area of the effusion hole, p and T denote the pressure and temperature of the ions inside the furnace. The equilibrium constant K is therefore given by the relation

$$K = \frac{p_{\mathrm{Sr}^+}p_e}{p_{\mathrm{Sr}}} = \frac{2\pi kT}{e^2S^2}\left(\frac{r^2+d^2}{r^2}\right)^2\frac{i_g^+ i_g}{p_{\mathrm{Sr}}}\sqrt{(m_{\mathrm{Sr}^+} m_e)}.$$

If all the pressures are expressed in atmospheres we have

$$K = \frac{p_{\mathrm{Sr}^+}p_e}{p_{\mathrm{Sr}}} = \frac{2\pi kT}{e^2S^2}\left(\frac{r^2+d^2}{r^2}\right)^2\frac{i_g^+ i_g^-}{p_{\mathrm{Sr}}}\frac{\sqrt{(m_{\mathrm{Sr}^+} m_e)}}{(1\cdot013\times10^6)^2}. \tag{1}$$

All the quantities occurring in this equation can be experimentally measured or are otherwise known, and hence K can be experimentally determined.

Now, according to Fowler's modification (1923) of the ionization formula, we have

$$\ln\frac{p_{\mathrm{Sr}^+}p_e}{p_{\mathrm{Sr}}} = -\frac{\chi_1}{kT}+\tfrac{5}{2}\ln T+\ln\frac{g_e(2\pi m_e)^{\frac{3}{2}}k^{\frac{5}{2}}}{h^3}-\ln b(T)+\ln b'(T), \tag{2}$$

where χ_1 denotes the energy required to ionize the unexcited atom (i.e. the ionization potential for the fundamental state), g_e the weight factor of the electron and is equal to 2, and $b(T)$ and $b'(T)$ are given by the relations

$$b(T) = g_1+g_2e^{-(\chi_1-\chi_2)/kT}+\ldots+g_ne^{-(\chi_1-\chi_n)/kT}+\ldots \tag{3}$$

and

$$b'(T) = g_{i,1}+g_{i,2}e^{-(\chi_{i,1}-\chi_{i,2})/kT}+\ldots+g_{i,n}e^{-(\chi_{i,1}-\chi_{i,n})/kT}+\ldots. \tag{4}$$

Here $g_{i,n}$ denotes the weight factor and $\chi_{i,n}$ the energy of the ionized atom in the nth quantum state.

Equation (2) can be written in the form

$$\log\frac{p_{\mathrm{Sr}^+}p_e}{p_{\mathrm{Sr}}} = -\frac{U}{4\cdot573T}+\tfrac{5}{2}\log T+\log\frac{(2\pi m_e)^{\frac{3}{2}}k^{\frac{5}{2}}}{h^3}+\log 2+\log b'(T)-\log b(T)$$

$$= -\frac{U}{4\cdot573T}+\tfrac{5}{2}\log T-6\cdot479+\log 2+\log b'(T)-\log b(T), \tag{5}$$

if the pressures are expressed in atmospheres, and the values of the other quantities are substituted. U is the energy of ionization of a gram-atom of strontium vapour in calories and is equal to $NeV_i/300J$, where V_i is the ionization potential of strontium, being equal to 5·667 V, and N is the Avogadro number. Thus $U = 130{\cdot}5$ kcal. For calculating $b(T)$ and $b'(T)$ we have to utilize the energy states of Sr and Sr^+ and their quantum weights. These are given by Bacher and Goudsmit (1932, pp. 448–55). A detailed calculation from formulae (3) and (4) shows that at the temperatures of the experiment the contribution of the excited states is negligible as they are not sufficiently developed on account of the high energy necessary for their excitation, and $b(T) = 1{\cdot}000$, $b'(T) = 2{\cdot}000$. Using these values in equation (5) we get the theoretical value of $\log K$. These are compared with the value of $\log K$ obtained experimentally from equation (1).

It will be observed that the quantity p_{Sr}, the pressure of strontium vapour in the main furnace, occurs in equation (1). There is considerable difficulty in evaluating this quantity. First, due to thermal transpiration this will be different from the vapour pressure of the solid strontium in the auxiliary furnace. As discussed by the author elsewhere (Srivastava 1939) the relation connecting these two cannot be exactly determined in this case. For our present purpose we shall assume the full Knudsen effect, i.e.

$$p/p' = \surd(T/T'),$$

where T and T' denote the temperatures of the main furnace and the auxiliary furnace. The error in $\log K$ due to this cause will, however, be negligible in comparison to the error caused by the possible errors in measuring T. The second difficulty arises from the fact that there is as yet no experimental determination of the vapour pressure of solid strontium. The only available data regarding the vapour pressure of strontium are those of Hartmann and Schneider (1929) in the liquid region. These have been utilized in the following manner for calculating the vapour pressure of solid strontium. The observed pressures are plotted on a graph in which $\log p$ represents the ordinate and $1/T$ the abscissa. The graph, as expected, is found to be a straight line (figure 1). This is extrapolated to the melting point of strontium for which the most probable value appears to be 1030° K as given by Weibke (1930). This gives the value $p = 1{\cdot}23$ mm. at $T = 1030°$ K. The same must be the vapour pressure over solid strontium at 1030° K as the melting point will differ very little from the triple point. This result has been utilized below in two different ways for calculating the vapour pressure of solid strontium at the temperatures prevailing in the experiment.

The first method consists in extending the above curve beyond the melting point by drawing a straight line whose slope is greater than the slope of the previous one by an amount equal to the latent heat of fusion of strontium (figure 1). Unfortunately the data regarding the various properties of strontium are very meagre. The latent heat of fusion of this substance does not appear to have been determined. In the circumstances the only available

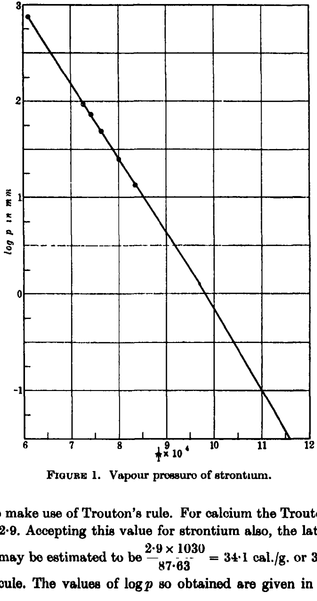


FIGURE 1. Vapour pressure of strontium.

course is to make use of Trouton's rule. For calcium the Trouton quotient ML/T_m is 2·9. Accepting this value for strontium also, the latent heat of strontium may be estimated to be $\frac{2\cdot 9\times 1030}{87\cdot 63} = 34\cdot 1$ cal./g. or 3·0 kcal. per gram-molecule. The values of $\log p$ so obtained are given in column (3) of table 1.

The second method consists in applying the vapour pressure formula

$$\log p\,(\text{atm.}) = -\frac{\lambda_0}{4\cdot573T} + \tfrac{5}{2}\log T - \frac{1}{4\cdot573}\int_0^T \frac{dT}{T^2}\int_0^T C_{p_s}\,dT - 1\cdot5885 + \tfrac{3}{2}\log M. \tag{6}$$

The only available data regarding the specific heat of solid strontium are those of Dewar (1913) which extend from − 253 to − 196° C. It is therefore impossible to evaluate accurately the specific heat integral in (6). The thermal expansion of strontium is also not known. The only available course is to assume that strontium obeys the Debye's formula for specific heat and on that assumption calculate the Debye temperature Θ which will make the observed specific heat agree with Debye's formula. In this way the value $\Theta = 102$ is obtained. This value has been utilized to evaluate the specific heat integral in (6) up to any temperature. Since, however, λ_0 is also not known, the extrapolated value of $\log p$ for the melting point as obtained above was utilized to give λ_0 in equation (6). The value of λ_0 thus obtained is 38,500 cal., a value which is quite probable in view of the fact that the latent heat of vaporization of liquid strontium at the melting point is 35·2 kcal. (Kelley 1935). This value of $\lambda_0 = 38{,}500$ cal. has been utilized in calculating the vapour pressure of solid strontium by means of equation (6). $\log p$, obtained in this way, is recorded in column (4) The mean of the corresponding readings in columns (3) and (4) is taken as the actual vapour pressure and is put in column (5). The small difference between the two sets of readings shows that the probability of error in the pressure thus calculated is small.

EXPERIMENTAL PROCEDURE

First the graphite tube was thoroughly degassed by prolonged intermittent heating for several hours after which the currents from the empty tube diminish to a negligible value. It was thus arranged that the currents during the experiment were much greater than those in the blank experiment. After the steady state had been attained in an experiment the currents were measured at 2 V with different currents through the electromagnet and for purposes of calculation the negative and positive currents without magnetic field were employed.

The temperature of the graphite furnace was measured by means of a Siemens' optical pyrometer and was correct to ± 10°. As the temperature of the furnace was not quite uniform the mean temperature of the tube was determined and employed for the purpose of calculation. In view of these uncertainties the slight difference of temperature between the inside

TABLE 1

Radius of limiting diaphragm = 4·2 mm.
Diameter of effusion hole = 1·99 mm.
Distance between hole and limiting diaphragm = 21·5 mm
Accelerating potential = 2 V.
Current sensitivity of the galvanometer = 1·53 × 10^{-9} amp./mm.

(1) Mean temperature of graphite furnace in °C	(2) Temperature of auxiliary furnace in °K	(3) log p_{Sr} in mm. from graph	(4) log p_{Sr} in mm. by calc.	(5) Mean value of log p_{Sr} in mm.	(6) Current through electromagnet in amp.	(7) Negative deflexion in mm.	(8) Positive deflexion in mm.	(9) log K (exp.) in atmospheres per cm.^{3}	(10) log K (theor.)	(11) U (exp.) in kcal.	(12) U (exp.) $- U$ (spect.)
1455	897	$\bar{2}$·875	$\bar{2}$·858	$\bar{2}$·867	0	330 × 3	9·5	$\overline{15}$·415	$\overline{15}$·707	132·7	+2·2
					6	10	9·5				
1465	906	$\bar{2}$·965	$\bar{2}$·950	$\bar{2}$·958	0	136 × 10	14	$\overline{15}$·634	$\overline{15}$·803	131·8	+1·3
					6	19	14				
1470	910	$\bar{1}$·01	$\bar{2}$·99	$\bar{1}$·00	0	172 × 10	15·5	$\overline{15}$·740	$\overline{15}$·856	131·4	+0·9
					6	22	15·5				
1495	916	$\bar{1}$·085	$\bar{1}$·053	$\bar{1}$·069	0	295 × 10	19·5	$\overline{14}$·009	$\overline{14}$·102	131·2	+0·7
					6	63	19·5				
1515	925	$\bar{1}$·165	$\bar{1}$·141	$\bar{1}$·153	0	139 × 30	30	$\overline{14}$·267	$\overline{14}$·304	130·7	+0·2
					6	44 × 3	30				
1465	906	$\bar{2}$·965	$\bar{2}$·950	$\bar{2}$·958	0	125 × 10	13	$\overline{15}$·565	$\overline{15}$·803	132·3	+1·8
					6	11	13				
1485	914	$\bar{1}$·05	$\bar{1}$·032	$\bar{1}$·041	0	180 × 10	19·5	$\overline{15}$·821	$\overline{14}$·006	131·9	+1·4
					6	17	19·5				
1510	920	$\bar{1}$·12	$\bar{1}$·092	$\bar{1}$·006	0	335 × 10	33	$\overline{14}$·359	$\overline{14}$·241	129·6	−0·9
					6	70	33				

and outside of the graphite tube as well as the slight difference between the true temperature and the black-body temperature of the graphite tube have been neglected.

The experiments were done with strontium supplied by Merck and the tube was freshly opened. The results obtained are recorded in table 1.

DISCUSSION OF RESULTS

It will be observed that the positive currents are much smaller than the negative currents. This is to be expected since the Sr^+ ion being $1{\cdot}61 \times 10^5$ times heavier than the electron, its rate of effusion will be about 400 times slower. The positive and negative currents will, however, not exactly follow this relation since their concentrations inside the furnace are not equal, for electrons are also produced by the graphite tube.

In view of the uncertainty in the determination of T by about 2 % as stated before, it is not proper to compare the value of K obtained experimentally, since a slight error in T produces a large error in K as $\log K$ varies as $1/T$. It is, however, reasonable to compare the values of $\log K$, as errors in this quantity will occur to the same percentage as the experimental errors in T. A comparison of the values of $\log K$ (column 9, table 1) obtained experimentally at different temperatures shows that $\log K$ increases with rise of temperature as demanded by the ionization theory. For the purpose of detailed comparison with the theory we have calculated $\log K$ theoretically from equation (5) for different temperatures and recorded the values in column (10) A comparison of columns (9) and (10) shows that the experimental and theoretical values of $\log K$ agree well within the limits of experimental error Another way of comparing these experimental results with the ionization theory is to calculate the values of U from the experimentally determined values of $\log K$ with the help of equation (5), and see how far they tally with the known spectroscopic value of 130·5 kcal. The experimentally determined values of U are recorded in column (11) and the difference between this value and the spectroscopic value $(U_{\text{exp}} - U_{\text{spect.}})$ is given in column (12). These differences are seen to be well within the limits of experimental error. Thus these experiments fully verify the ionization theory with reference to strontium.

In conclusion I wish to express my sincere thanks to Professor M. N. Saha, F.R.S., for his kind interest in the work. My thanks are also due to the Royal Society of London for a grant which covered the cost of the furnace and its accessories.

SUMMARY

In this paper the thermal ionization of strontium vapour has been experimentally investigated by an apparatus already described by the author elsewhere in his work on barium. Experiments have been carried out at various temperatures and pressures of strontium vapour and the equilibrium concentrations of Sr^+ and electrons inside the furnace have been obtained by allowing them to effuse out through a narrow opening. From these the equilibrium constant and the energy of ionization have been calculated. The results obtained agree, within the limits of experimental error, with the theory of thermal ionization and the known spectroscopic value of the ionization potential of strontium.

REFERENCES

Bacher, R. F. and Goudsmit, S. 1932 *Atomic energy states*, 1st ed New York and London. McGraw-Hill Book Co.
Dewar, J. 1913 *Proc. Roy. Soc.* A, **89**, 158.
Fowler, R. H. 1923 *Phil. Mag.* **45**, 1.
Hartmann, H. and Schneider, R. 1929 *Z. anorg. Chem.* **180**, 275.
Kelley, K. K. 1935 *Bull U.S. Bur. Mines*, no **383**
Kilian, T. J. 1926 *Phys. Rev* **27**, 578.
Langmuir, I. and Kingdon, K H 1925 *Proc. Roy Soc* A, **107**, 61
Meyer, E. 1930 *Ann. Phys., Lpz* , **4**, 357.
Morgulis, N. 1934 *Z. Sowjet.* **5**, 221.
Saha, M N. 1920*a* *Phil Mag.* **40**, 472.
— 1920*b* *Phil. Mag* **40**, 809.
— 1921 *Proc. Roy. Soc.* A, **99**, 135.
Saha, Sur and Majumdar 1927 *Z. Phys.* **40**, 648
Saha, M. N. and Tandon, A. N. 1936 *Proc. Nat Acad Sci. India*, **6**, 212.
Srivastava, B. N 1939 *Proc. Nat. Inst. Sci. India*, **5**, 316
— 1940 *Proc. Roy. Soc* A, **175**, 26.
Weibke 1930 *Z anorg Chem* **193**, 302.

The kinetics of mutarotation in solution

By J. C. Kendrew and E. A. Moelwyn-Hughes

The University Chemical Laboratory, Cambridge

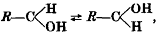
(*Communicated by R. G. W. Norrish, F.R.S.—Received* 14 *May* 1940)

The kinetics of the mutarotation of representative reducing sugars from the pentose, hexose and disaccharide series have been investigated polarimetrically over wide temperature ranges in aqueous solution.

The dependence of the velocity coefficient, k, upon temperature is fairly well reproduced by an equation of the form $\ln k = C + (J/R)\ln T - E/RT$. The true energy of activation, E, is found to be some 6000 calories greater than the apparent value afforded by the Arrhenius equation at room temperature. J/R has a value of -10, which is identified as the number of oscillators contributing to the activation.

The constants C, J and E of this equation are discussed, with reference to many reactions, in terms of a theory of unimolecular reactions in solution.

The mutarotation of reducing sugars, which may be represented by the equation



$$R{-}C\begin{matrix}\diagup\,H\\ \diagdown\,OH\end{matrix} \rightleftharpoons R{-}C\begin{matrix}\diagup\,OH\\ \diagdown\,H\end{matrix}, \qquad (1)$$

has long been recognized as one of the simplest of chemical changes. In this paper we present evidence that the number of atoms taking part in the activation necessary for conversion is probably not more than seven, and not less than four. The distribution of the critical energy is thus highly localized, and cannot spread from the potentially aldehydic group farther than to one neighbouring water molecule or to the contiguous carbon atom in the pyranose or furanose ring, RC. This conclusion is borne out by the new data which we have obtained from the polarimetric examination of some representative sugars taken from the pentose, hexose and disaccharide series.

The dependence upon temperature, T, of the velocity coefficient, k, governing the mutarotations is approximately reproduced by the equation

$$\ln k = C + \frac{J}{R}\ln T - \frac{E}{RT}, \qquad (2)$$

in which the constants, C, J and E, are specific for each sugar. Mutarotations thus come into line with hydrogen-ion catalysis (Moelwyn-Hughes 1934), ionization (Moelwyn-Hughes 1938*a*), hydrolysis (Moelwyn-Hughes 1938*b*) and decarboxylation (Johnson and Moelwyn-Hughes 1940) as instances of the inapplicability of the Arrhenius equation. These changes are so diverse that we may now venture on the generalization that the apparent energy of activation of all unimolecular reactions in aqueous solution decreases with a rise in temperature. The explanation of this experimental rule must be that the critical energy is used in part to increase the internal energy of the reacting solute molecule, and in part to release the solute molecule from its imprisonment by solvent neighbours.

EXPERIMENTAL METHOD

The mutarotations were investigated polarimetrically, by measuring, from time to time, the specific rotation of an aqueous solution which had been prepared by weight and pre-heated. The verniers of the polarimeter employed—a sensitive instrument constructed to the specifications of the late Professor Lowry—could be read to 0·002°. Our readings were generally reproducible only to 0·01°, and refer to light from the sodium vapour lamp. The half-shadow angle was maintained constant at 6·5°. Preliminary experiments performed on a number of sugars, using an ordinary water-jacketed 4 dm. tube, enabled us immediately to reproduce the most reliable values in the literature for the rate of mutarotation. In our attempt to increase the accuracy of the already accurate information available, we succeeded in tracing the greatest uncertainty to fluctuations of temperature in the jacketed tubes.

Water was pumped through an ordinary metal jacket from a thermostat which was capable of maintaining a temperature steady to 0·02°, and the temperature of the circulating water was measured by passing it through lagged bottles on entering and leaving the tube. These bottles were fitted with calibrated thermometers which could be read to 0·01°. When the temperature of the water differed by more than a few degrees from that of the room, a temperature gradient was detected along the tube, and could not be eliminated by lagging. The gradient amounted to as much as 0·04° per dm. at 313° K.

To overcome this difficulty, a special type of polarimeter tube was constructed, according to the plan shown in figure 1. The circulating water flows first along the inner jacket, and back along the outer one. The inner jacket is then losing heat, not to the surroundings, but to the outer jacket, the temperature of which is only slightly lower than that of the inner one. Experiments with thermometers arranged inside the glass tube showed that this device of compensating temperature gradients ensured uniformity of heating under all the conditions of the present investigation. The most satisfactory method of measuring the temperature of the solution contained in the glass tube is to include a flooded, rubber-bunged thermometer bottle in the system, as near as possible to the inflow to the metal tube. There is no detectable difference between the temperature registered on a thermometer so fixed and the temperature of the solution inside the glass tube.

The plain glass polarimeter tubes used to contain the solution were fixed into the jacket by rubber washers, as indicated in the diagram, and were removed completely from the metal container after each run.

A stream of air playing on the windows prevented dew from obscuring vision at low temperatures.

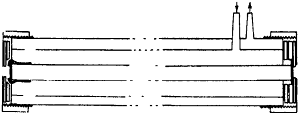


FIGURE 1 Glass polarimeter tube enclosed in concentric metal jackets to ensure constancy of temperature along its length.

In carrying out an experiment, about 3 g. of dried sugar were weighed in a calibrated 50 ml flask, which, along with a second flask containing pure water, was placed in the thermostat. Water was meanwhile circulating round the empty polarimeter tube. The solution was made at a known instant. While keeping the metal jackets full of thermostat water, the glass tube was removed, filled with the solution, and the apparatus was reassembled. The time taken by these operations was such that the first optical reading could be taken some 4 min. after the reaction had started. In general, however, we place little reliance on readings taken during the first 10 min. Each accepted rotation is the mean of three readings straddled evenly about the corresponding time.

Purity of materials

The solvent employed throughout was water prepared for electrical conductivity purposes by distillation in an all-quartz still. With the exception of glucose, the sugars were supplied as bacteriologically pure, and we were unable to increase their purity by crystallization. In the case of the monohydrate of α-lactose, for example, the material supplied gave an initial specific rotation of $+83\cdot37°$, a final specific rotation of $+52\cdot44°$, and an observed velocity coefficient of mutarotation of $1\cdot843 \times 10^{-4}$ sec.$^{-1}$ at $293\cdot19°$ K. After recrystallization, the figures were $+83\cdot35$, $+52\cdot31$, and

$1{\cdot}811 \times 10^{-4}$ at 293·18° K. The two sets of values agree well with each other and with the accepted constants (Hudson 1903). The specimen of α-glucose supplied as an analytical reagent compared very favourably with the pure sugar which we prepared according to instructions kindly given us by Professor W. N. Haworth, F.R.S. The specific rotatory power of initial and equilibrated solutions are given in table 1, the bracketed figures in which are due mainly to Hudson and Yanovsky (1917) and in part to Moelwyn-Hughes (1928).

TABLE 1. SPECIFIC ROTATIONS OF INITIAL AND EQUILIBRATED SOLUTIONS OF SUGARS

Sugar	$[\alpha_0]^{293\cdot1^\circ\,\mathrm{K}}_{5893\,\mathrm{A}}$	$[\alpha_\infty]^{293\cdot1^\circ\,\mathrm{K}}_{5893\,\mathrm{A}}$
α-Arabinose	+182·6 (+175)	+103·85 (+105·0)
α-Xylose	+92·8 (+92·0)	+18·92 (+19·0)
β-Mannose	−16·44 (−17)	+14·43 (+14·5)
α-Glucose	+110·0 (+113·4)	+52·56 (+52·56)
α-Lactose monohydrate	+83·4 (+85·5)	+52·44 (+52·54)

Formulation of the kinetics and statics of mutarotation

Four out of the five sugars examined were converted from the α- to the β-form according to the scheme of opposing unimolecular reactions (Lowry 1899), which, in the special case of a system containing initially only one modification, may be represented as follows:

$$\begin{array}{ccc} A & \underset{k_\beta}{\overset{k_\alpha}{\rightleftarrows}} & B. \\ a-x & & x \end{array}$$

If the initial concentration, a, has been reduced to $a-x$ in t sec., the rate of conversion obeys the differential law

$$\frac{dx}{dt} = k_\alpha(a-x) - k_\beta x, \tag{3}$$

which gives on integration

$$x = a\frac{k_\alpha}{k}(1-e^{-kt}), \tag{4}$$

where

$$k = k_\alpha + k_\beta. \tag{5}$$

The equilibrium relations

$$\frac{k_\alpha}{k_\beta} = \frac{x_\infty}{a-x_\infty} = K, \tag{6}$$

follow from the equivalent suppositions of zero rate (equation (3)) or infinite time (equation (4)).

In the study of reversible reactions, it is expedient to determine the ratio of the velocity coefficients before evaluating either.

Static measurements

In the third column of table 2 are to be found the initial specific rotations of the pure sugars, and in the fourth column the specific rotations of the equilibrium mixtures. Each pair of figures refers to solutions of approximately 0·25 M concentration, and to the temperature recorded in column 2. The infinity rotations were obtained by retaining the solution at the experimental temperature until further rotational change could not be detected, they are regarded as accurate to within ± 0·02°. The values for mannose, which was the first sugar to be examined by us, and the bracketed values for other sugars, refer to systems which were allowed to react to completion at room temperature. Independent experiments with mannose showed $[\alpha_\infty]$ to fall from 14·14 to 13·92° as T is raised from 290·22 to 313·25° K. The initial rotations were obtained by logarithmic extrapolation to zero time, and are limited in accuracy by the rate of dissolution of the sugar and by the temperature irregularities attending the charging of the tubes. The error due to the former effect is specific, being larger, for example, in the case of lactose than in the case of glucose. The greater error is due to the latter effect, and increases with the difference between the experimental and room temperatures When this difference is positive, $[\alpha_0]$ is too high, and when the difference is negative, $[\alpha_0]$ is too low. On plotting the experimental values of the initial specific rotations as a function of T, we should therefore obtain a point of inflexion at room temperature. This behaviour, as the figures in table 2 reveal, is observed for all the sugars. As may further be seen from these data, a cusp in the curve appears at low temperatures, and a rapid increase of $[\alpha_0]$ as T is raised. Without adopting a much more elaborate experimental technique, it is impossible to disentangle the errors of determination from the genuine dependence of the initial rotatory power upon temperature. Such a dependence we consider to be a real one in all the cases examined, though its magnitude is less than in the pronounced case of fructose.

In order to estimate the ratio of the velocity coefficients of the two transformations, we provisionally regard the initial specific rotations, $[\alpha_0^a]$ and $[\alpha_0^b]$, of the α- and β-modifications to be independent of temperature. In terms of these constants, and of the specific rotation $[\alpha_\infty]$ of the equilibrated solution, equation (6) becomes

$$K = \frac{[\alpha_\infty] - [\alpha_0^a]}{[\alpha_0^b] - [\alpha_\infty]}. \tag{7}$$

For α-xylose, β-mannose, α-glucose and α-lactose, we employ the values obtained by us at 298·1° (table 1). The accepted figures for β-xylose (−20°), α-mannose (+34°), β-glucose (+19·7°) and β-lactose (+33·26°) are due to Hudson and Yanovsky (1917). Table 4 contains the experimental values of the equilibrium constants, as well as those reproduced by the equation

$$K = e^{-\Delta G^0/RT} = e^{\Delta S^0/R}e^{-\Delta H^0/RT}. \tag{8}$$

TABLE 2. STATIC ROTATIONAL DATA

Sugar	T° K	$[\alpha_0]_{5893\,\text{A}}$	$[\alpha_\infty]_{5893\,\text{A}}$	K_{observed}	K_{formula}
α-Xylose	273·45	+91·52	+(18·81)	—	2·019
	278·03	91·07	(18·82)	—	1·986
	283·03	89·10	(18·85)	—	1·951
	287·96	91·68	18·45	1·921	1·920
	293·06	92·81	18·92	1·886	1·887
	298·09	94·39	19·26	1·860	1·858
	303·17	101·6	19·72	1·827	1·821
	308·17	110·3	20·05	1·804	1·803
β-Mannose	273·41	−16·41	+13·95	—	
	278·05	16·34	13·75		
	283·09	16·11	13·94		
	288·09	15·78	14·02		
	293·13	16·44	14·03	0·656	0·660
	298·12	16·68	14·00	0·657	
	303·06	17·60	13·85	0·665	
	308·12	19·06	13·95	0·660	
α-Glucose	273·32	+109·8	+52·52	0·571	0·575
	278·26	109·2	52·52	0·571	
	283·12	107·4	52·70	0·574	
	288·15	108·6	52·61	0·572	
	293·09	108·8	52·53	0·571	
	298·06	110·3	52·69	0·574	
	303·07	112·4	52·94	0·577	
	308·16	115·2	52·99	0·579	
	313·14	121·3	52·90	0·577	
	318·19	127·1	52·95	0·579	
α-Lactose	288·04	+85·09	+(53·08)	—	1·583
	293·19	83·37	52·44	1·619	1·618
	298·11	84·84	52·23	1·647	1·650
	303·28	84·59	52·16	1·657	1·685
	308·18	84·98	51·64	1·732	1·717
	313·21	86·28	51·63	1·734	1·750
	318·12	85·96	51·34	1·777	1·781
	322·23	88·63	51·14	1·809	1·807

The standard increases in thermodynamic potential (G), entropy (S) and heat content (H) associated with the conversion of a gram-molecule of α-sugar into β-sugar are summarized in table 3. A comparison of ΔS^0 for the mutarotation of α-lactose monohydrate with the molar entropy of fusion of water (5·3 cal./deg.) lends support to Hudson's view of the mechanism (Hudson 1903), but we incline to the opinion that the dehydration merely plays a part in, but does not constitute, the isomerization. There is no necessity to introduce the conception of dehydration into the other cases under discussion. The molar entropy of the β-forms of xylose, mannose and glucose is roughly $0{\cdot}43R$ less than the molar entropy of the α-forms, and suggests a very simple configurational rearrangement which must resemble a local freezing.

TABLE 3. THERMODYNAMIC MAGNITUDES FOR MUTAROTATION EQUILIBRIA IN AQUEOUS SOLUTION

Sugar	$K_{298{\cdot}1}$	$\Delta G^0_{298{\cdot}1}$ cal /g.mol.	$\Delta H^0_{298{\cdot}1}$ cal./g.mol.	$\Delta S^0_{298{\cdot}1}$ cal /g.mol. deg.
Xylose	1·86	−370	−550	−0·61
Mannose	0·66	+250	0	−0·84
Glucose	0·575	+330	0	−1·10
Lactose	1·65	−300	+710	+3·4

Kinetic measurements

With the exception of arabinose, the sugars investigated have been found to obey the law of reversible unimolecular changes. In terms of the experimental rotation, α, equation (4) becomes

$$\alpha_t - \alpha_\infty = (\alpha_0 - \alpha_\infty)\, e^{-kt}, \tag{9}$$

in which k, the observed unimolecular constant, is the sum of the constants for the direct and inverse reactions (equation (5)). Representative examples of the applicability of this equation to the mutarotation of xylose, mannose,

TABLE 4. α-XYLOSE

$a = 0{\cdot}4065$ g.mol./l. $T = 303{\cdot}07°$ K. $k = 2{\cdot}087 \times 10^{-4}$ sec.$^{-1}$

	$\alpha_t - \alpha_\infty$	
t (min.)	Obs.	Calc.
5	10·51	10·64
8	7·30	7·40
11	4·93	5·02
14	3·43	3·44
17	2·36	2·37
20	1·64	1·63
23	1·11	1·12
26	0·77	0·77
29	0·53	0·53

glucose, and lactose are shown in tables 4–7. The precision and degree of reproducibility of the velocity coefficients increase with the total observable change in rotation. Examples are furnished in table 8.

TABLE 5. β-MANNOSE

$b = 0\cdot2939$ g.mol./l. $T = 293\cdot12°$ K. $k = 7\cdot06 \times 10^{-4}$ sec.$^{-1}$

	$\alpha_t - \alpha_\infty$	
t (min.)	Obs.	Calc.
15	3·34	3·34
20	2·70	2·70
25	2·20	2·19
30	1·77	1·77
35	1·43	1·43
40	1·16	1·16
45	0·94	0·94

TABLE 6. α-GLUCOSE

$a = 0\cdot3193$ g.mol./l. $T = 278\cdot26°$ K. $k = 5\cdot13 \times 10^{-5}$ sec.$^{-1}$

	$\alpha_t - \alpha_\infty$	
t (min.)	Obs.	Calc.
20	12·23	12·26
30	11·86	11·88
40	11·50	11·52
50	11·17	11·17
60	10·84	10·83
75	10·35	10·35
90	9·88	9·88
105	9·43	9·43
125	8·87	8·87
154	8·10	8·10
161	7·90	7·94
180	7·49	7·49
190	7·27	7·26
210	6·82	6·83
220	6·61	6·62
240	6·22	6·23
260	5·83	5·85
280	5·49	5·50
300	5·16	5·17
330	4·72	4·72
360	4·28	4·30
390	3·90	3·92
420	3·55	3·58
450	3·23	3·26
480	2·94	2·97

TABLE 7. α-LACTOSE

$a = 0{\cdot}1625$ g.mol./l. $T = 322{\cdot}23^\circ$ K. $k = 2{\cdot}49 \times 10^{-3}$ sec.$^{-1}$

t (min.)	$\alpha_t - \alpha_\infty$	
	Obs.	Calc.
5	4·12	4·15
7	3·08	3·08
9	2·28	2·29
11	1·70	1·69
13	1·25	1·26
15	0·95	0·93
17	0·69	0·69
19	0·51	0·51

TABLE 8. REPRODUCIBILITY OF KINETIC RESULTS

Sugar	Temp °K	Molar concentration	$k \times 10^4$ sec.$^{-1}$	Average value
Mannose	293·13	0·2438	7·218	7·166
		0·2939	7·060	
		0·3635	7·219	
Glucose	293·13	0·2590	2·479	2·473
		0·2665	2·470	
		0·2802	2·471	
		0·2914	2·471	
Lactose	303·28	0·2836	4·899	4·883
		0·2904	4·873	
		0·3272	4·877	

The influence of temperature

Before recording our results on the temperature dependence of the velocity coefficient, we shall resolve the composite constants into those governing the individual unimolecular processes. According to equations (5) and (6), we have

$$k_\alpha = k\frac{K}{K+1}, \tag{10}$$

and

$$k_\beta = k\frac{1}{K+1}. \tag{11}$$

It follows from the definitions

$$E_{A,\alpha} = RT^2\left(\frac{d\ln k_\alpha}{dT}\right)_P; \quad E_{A,\beta} = RT^2\left(\frac{d\ln k_\beta}{dT}\right)_P; \quad E_A = RT^2\left(\frac{d\ln k}{dT}\right)_P, \tag{12}$$

that

$$E_{A,\alpha} = E_A + \Delta H\frac{1}{K+1}, \tag{13}$$

and that

$$E_{A,\beta} = E_A - \Delta H\frac{K}{K+1}. \tag{14}$$

All the results for the velocity coefficients and the apparent energies of activation summarized below refer to the rate of conversion of the α-modification into the β-modification. In the case of each sugar, the apparent energies of activation afforded by equation (12) are found to decrease as the temperature is raised. The velocities and energies measured at 298·1° K are summarized in table 9, in the last column of which we give the rate of decrease of the activation energy and the probable error. The latter figures have been obtained by considering sets of data in limited temperature regions wherein the energy of activation can be measured to within about 20 cal. Thus, for example, in a set of experiments made with mannose, we found $E_{A,\alpha}$ to be $16{,}986 \pm 13$ cal at 278·25° K, and $16{,}562 \pm 23$ cal. at 299·35° K. Considered alone, these figures would give us $J = -20{\cdot}1 \pm 1{\cdot}7$ cal./g.mol. deg. On account of the uncertainty which attaches to the exact values of J for the different sugars, we have taken a mean value of 20·5 cal./g.mol. deg. to hold for them all. The expressions for the variation of the velocity coefficient with temperature then become:

TABLE 9. KINETIC DATA FOR THE CONVERSION OF α-SUGARS AT 298·1° K

Sugar	$k_\alpha \times 10^{-4}$ sec.$^{-1}$	$E_{A,\alpha}$ cal./g mol.	$J = \left(\frac{dE_{A,\alpha}}{dT}\right)_p$
Xylose	8·578	16,245	$-20{\cdot}4 \pm 8{\cdot}4$
Mannose	4·572	16,375	$-20{\cdot}5 \pm 2{\cdot}8$
Glucose	1·463	16,945	$-19{\cdot}6 \pm 5{\cdot}9$
Lactose	1·869	17,225	$-21{\cdot}9 \pm 5{\cdot}6$

$$\begin{aligned}
&\text{For xylose:} && \ln_e k_\alpha = 89{\cdot}5367 - (20{\cdot}5/R)\ln_e T - 22{,}355/RT,\\
&\text{For mannose:} && \ln_e k_\alpha = 89{\cdot}0548 - (20{\cdot}5/R)\ln_e T - 22{,}485/RT,\\
&\text{For glucose:} && \ln_e k_\alpha = 88{\cdot}8789 - (20{\cdot}5/R)\ln_e T - 23{,}055/RT,\\
&\text{For lactose.} && \ln_e k_\alpha = 89{\cdot}5935 - (20{\cdot}5/R)\ln_e T - 23{,}335/RT.
\end{aligned}$$

The velocity coefficients reproduced by these formulae are compared in tables 10–13 with the experimental values. We have taken R as 1·9869 cal./g.mol. deg., and $T° \text{K} = 273{\cdot}1 + t° \text{C}$.

*A comparison of the experimental values with those reproduced by equation (2) and by the Arrhenius equation ($J = 0$) is made in table 14. The calculated values of the velocity coefficients are denoted, respectively, by l and m. *Assuming that J is zero, we find, by the method of least squares,* that the best average value of the energy of activation for the whole temperature region examined is 16,950 cal. If now we define e_1 as $(l - k_\alpha)/100$, then Σe_1^2 becomes 18 in round figures. When $E_A = 16{,}900$, $\Sigma e_1^2 = 19$, and when $E_A = 17{,}000$, $\Sigma e_1^2 = 23$. By the same method, a comparison of the experimental results with equation (2) fixes $-J$ within the limits 20 ± 3. The best fit is found when J is -18, but, for the reasons given earlier, a rounded

---

* Paragraph added 16 July 1940.

TABLE 10. THE MUTAROTATION OF α-XYLOSE IN WATER

	$k_\alpha \times 10^5$	
T° K	Observed	By formula
273·45	6 94	6 94
278 03	11 6	11·5
283 03	19 7	19 6
287·96	32 9	32·3
293·06	52·4	52·3
298·09	84·3	85·5
303·17	134 9	135·2
308·17	207 7	208 4

TABLE 11. THE MUTAROTATION OF α-MANNOSE IN WATER

	$k_\alpha \times 10^5$	
T° K	Observed	By formula
273·41	3 610	3 620
278 05	6·115	6 054
283 09	10·68	10·43
288·09	17 53	17·38
293 13	28 49	28 57
298·12	45 52	45·81
303·06	71 76	71·76
308 12	111·5	111·7

TABLE 12 THE MUTAROTATION OF α-GLUCOSE IN WATER

	$k_\alpha \times 10^5$	
T° K	Observed	By formula
273·32	1 052	1·051
278·26	1·864	1·857
283 12	3·195	3 177
288·15	5·451	5 417
293·13	9·028	8·995
298·06	14 36	14·57
303·07	23·35	23 36
308·16	36·81	37·02
313 14	56 82	57·11
318·19	87·38	87·20
323 13	129·6	129·9

TABLE 13. THE MUTAROTATION OF α-LACTOSE IN WATER

T° K	$k_\alpha \times 10^5$ Observed	$k_\alpha \times 10^5$ By formula
288·04	6·77	6·73
293·19	11·48	11·48
298·11	18·79	18·72
303·28	30·65	30·68
308·18	48·19	48·14
313·21	74·96	75 11
318 12	113·8	114·4
322·23	160·3	160 1

value for all the sugars has been accepted in the summarizing tables. The quantities l and m of table 14 are given by the relations

$$\log_{10} l = 8{\cdot}5983 - 16{,}950/4{\cdot}571T$$

and

$$\log_{10} m = 34{\cdot}981 - (18/R)\log_{10} T - 22{,}317{\cdot}6/4{\cdot}571T.$$

TABLE 14. A COMPARISON OF THE EXPERIMENTAL RESULTS FOR GLUCOSE WITH EQUATION (2), TAKING J AS 0 AND AS -18

T° K	$k_\alpha \times 10^5$	$l \times 10^5$	e_1	e_1^2	$m \times 10^5$	e_2	e_2^2
273·32	1 052	1·074	+2	+4	1·052	0	0
278·26	1·874	1·871	0	0	1·856	−1	1
283·12	3·195	3 168	−1	1	3 173	−2/3	1/2
288·15	5 451	5·363	−7/4	3	5 408	−4/5	2/3
293·09	9 020	8·837	−2	4	8 948	−4/5	2/3
298·06	14·36	14·36	0	0	14·56	+4/3	7/4
303 07	23 35	23 07	−5/4	3/2	23·35	0	0
308·16	36·64	36·73	−1/4	0	37·05	+1	1
313·14	56·82	57·07	+1/2	1/4	57·24	+3/4	1/2
318·19	87·38	87 98	+1/2	1/4	87 54	+1/5	0
323·13	129·6	132 6	+2	4	130 9	+1	1
				$\Sigma e_1^2 = 18$			$\Sigma e_2^2 = 7$

The percentage error, e_2, according to the second of these equations is $(m-k)/100$. It is seen from the table that the equation of Arrhenius reproduces the experimental facts to a very satisfactory first approximation. The agreement is best in the medium temperature interval, to which the apparent energy truly belongs. The best fit of the Arrhenius equation, however, always fails, in reactions of this kind, to account with sufficient accuracy for the velocity at the extreme temperatures. It is also appreciated from table 14 that equation (2) is in better conformity with our results; e_2 is not only smaller than e_1 but is more evenly spread. The refined analysis of the temperature dependence of the velocity coefficient of reactions in solution is more fully dealt with in a concurrent publication on the decomposition of phenyldiazonium chloride (Johnson and Moelwyn-Hughes 1940).

Comparison with earlier work

The velocity of mutarotation of glucose in water at 293·10° K is a well-known kinetic standard, assessed by the classical work of Riiber (1923) and of Richards, Lowry and Faulkner (1927) as $(2{\cdot}47 \pm 0{\cdot}03) \times 10^{-4}$ sec.$^{-1}$. The value obtained in this work is $2{\cdot}475 \times 10^{-4}$.

The apparent energy of activation (E_A) which may be obtained by analysing the data of Hudson and Sawyer on the mutarotation of mannose at 293·35° is 16,460 ± 320 cal. The corresponding value yielded by equation (13) at this temperature is 16,470 The directly observed values at 295·63°, 300·59° and 305·59° are 16,480, 16,550 and 16,540 respectively, giving a mean of 16,525 ± 45 cal

In the case of lactose, we cannot assess E_A closer than 110 cal. at 294·17°. Our value of 17,033 agrees with, and is seen to be more accurate than, that derived from the results of Hudson (1903), which is 16,880 ± 670 at the same temperature.

In the case of glucose, the familiar work of Hudson and Dale (1917) has long been regarded as a good instance of the applicability of the Arrhenius equation. With one exception, however, their E_A values fall regularly as T is raised. This fact was first noted by G. F. Smith and M. C. Smith (1937), whose own experiments confirm their deduction and yield an approximate value of $J = -29$.

The mutarotation of arabinose

The changes in optical rotation observed during the mutarotation of pure α-arabinose in aqueous solution do not obey the law of reversible unimolecular reactions. The velocity coefficient calculated by means of the equation

$$k = \frac{1}{t_2 - t_1} \ln \frac{\alpha_1 - \alpha_\infty}{\alpha_2 - \alpha_\infty}, \tag{15}$$

consequently varies with the extent of chemical change, as instanced in table 15. Approximate values for the composite velocity constants at zero and infinity times are summarized in table 16, along with the corresponding values of k_α. These results indicate the existence of more than two modifications of arabinose in solution; and until these have been characterized by rotatory constants, the kinetic interpretation is impossible.

TABLE 15. α-ARABINOSE

$a = 0{\cdot}3434$ g mol./l. $T = 273{\cdot}58°$ K

$\frac{1}{2}(t_1+t_2)$ min.	$k \times 10^4$ sec.$^{-1}$
15	1·681
25	1·581
35	1·551
45	1·546
55	1·486
65	1·435
75	1·412
85	1·409
110	1·401
130	1·329
150	1·349

TABLE 16. THE MUTAROTATION OF α-ARABINOSE IN WATER

T	K	$k_0 \times 10^4$	$k_\infty \times 10^4$	$k_x \times 10^4$
273·58	1·279	1·79	1·34	0·81
278·25	1·349	2·95	2·47	1·43
282·98	1·422	4·88	4·16	2·43
287·97	1·500	7·75	6·96	4·30
293·08	1·581	12·1	11·4	6·90
298·03	1·660	18·9	17·9	11·0
303·21	1·742	29·4	27·5	17·5

DISCUSSION

The kinetics of reactions in solution may now be regarded as in the third stage of development. The first stage coincides with the elucidation of molecular mechanism as revealed by measurements made at constant temperature. The dependence of the velocity of reaction on the concentration of reactants leads to the order of reaction; its dependence on added substances to the detection of catalysis; and its dependence on salt concentration and dielectric capacity of the medium to the magnitude of the electrostatic disturbances. The second stage is represented by the approximate determination of the variation of the velocity coefficient with temperature. This phase is marked by the study of the equation $\ln k = A - B/T$, which has the form so often assumed by simple physicochemical systems which obey the Boltzmann statistics. The first coherent presentation of the whole available data has, in fact, centred round a comparison of the constant A with certain theoretical frequencies of vibrations within the molecule in the case of unimolecular reactions, and of encounters between molecules in bimolecular and termolecular reactions. The comparison rests on the necessary assumption that RB is the energy of activation. The third stage, to which the present series of investigations belongs, is concerned with the modes of allocation of the critical energy among the oscillators which become excited or damped during activation. Its development requires data which are both more extensive and precise than those ordinarily available, and an abandonment of the classical for quantal statistics.

As indicated in the opening sections of this paper, we now consider the experimental evidence sufficiently strong to warrant the generalization that unimolecular reactions in aqueous solution conform more closely to the following equation (cf. equation (2))

$$\log_{10} k = c - b \log_{10} T - a/T, \tag{16}$$

than to the equation of Arrhenius. According to classical and quantal theories, the integer b is the number of feeble oscillations which take part in the activation. For example, the unimolecular constant, k_α, governing the conversion of α-glucose into β-glucose, may be expressed as follows:

$$\log_{10} k_\alpha = 37{\cdot}814 - 10 \log_{10} T - 5039{\cdot}34/T. \tag{17}$$

Thus ten oscillations are involved, though the number is uncertain between 7 and 13 (see table 9). In the hydrolysis of methyl bromide, the number is so much higher (34 ± 4) that we may rule out a hydrolytic mechanism for the mutarotation, which may be legitimately considered as a highly localized activation confined to the potentially aldehydic group and possibly one water molecule. This relative simplicity in the mechanism, coupled with the similarity of the outer sugar structure to water, suggests that our observations at constant pressure are not very different from those which would be obtained at constant volume. We shall therefore compare our results directly with equation (12) of our preceding paper (Johnson and Moelwyn-Hughes 1940):

$$k_\alpha = \nu \times \frac{\prod^{r} (1 - e^{-h\nu_r/kT})^{-1} \prod^{s} (kT/h\nu_s)}{\prod^{p} (1 - e^{-h\nu_p/kT})^{-1} \prod^{q} (kT/h\nu_q)} e^{-\epsilon/kT}. \tag{18}$$

The effective mean frequency, defined by the relation

$$\bar{\nu} = \left[\nu \times \frac{\nu_1 \nu_2 \ldots \nu_q}{\nu_1' \nu_2' \ldots \nu_s}\right]^{(q-s+1)^{-1}}, \tag{19}$$

is then found to be $6{\cdot}61 \times 10^{12}$ sec.$^{-1}$, which lies near to the frequency kT/h ($6{\cdot}24 \times 10^{12}$ at $T = 298{\cdot}2°$ K). The corresponding mean wave number is 220 cm.$^{-1}$. These results support the supposition of activation among feeble oscillators, many of which would appear to be of the hindered rotational type (Fox and Martin 1940).

The exact comparison of theory and experiment requires information on the complicated question of the dependence of many of these oscillatory frequencies on temperature. Until such information appears, we may regard the present theory as sufficiently clear in outline though blurred in detail.

In table 17 we give the results of similar calculations for other reactions studied in this series. Columns 2 and 3 contain the experimental values of the energy of activation and effective mean frequency. In the last column is found the number of quanta, $h\bar{\nu}$, which would fully account for E_0. We infer that the activation energy in mutarotation is not drawn solely from oscillatory motions or that each oscillator contributes many quanta. In the other cases, the figures in columns 4 and 5 tally. $q-s$ is about 21 for trinitrobenzoic acid and about 9 for trichloracetic acid; one is naturally inclined to attribute the difference of 12 to the three nitro groups, which may claim 4 each.

TABLE 17

Reactant in aqueous solution	E_0 cal./g.mol	$\bar{\nu} \times 10^{-13}$ sec $^{-1}$	$q-s$ obs.	$E_0/h\bar{\nu}$
α-Glucose	23,055	0·66	10 ± 3	36·7
Tribromacetate ion	39,610	5·25	10 ± 5	8·0
Trichloracetate ion	42,910	5·70	10 ± 5	7·8
Methyl bromide	46,820	1·66	34 ± 4	29·7
2 : 4 : 6-Trinitrobenzoate ion	52,300	2·95	23 ± 5	18·7

We are greatly indebted to Professor W. N. Haworth, F.R.S., for advice in the purification of glucose, to the Chemical Society for a grant which partly defrayed the cost of chemicals, and to the Department of Scientific and Industrial Research for a maintenance allowance. Mr P. Johnson gave us most valuable assistance in completing our programme of work, which would otherwise have been left unfinished.

REFERENCES

Fox and Martin 1940 *Proc. Roy. Soc.* A, **174**, 234.
Hudson 1903 *Z. phys. Chem.* **44**, 487.
Hudson and Dale 1917 *J. Amer. Chem. Soc* **39**, 320
Hudson and Sawyer 1917 *J. Amer. Chem. Soc.* **39**, 470
Hudson and Yanovsky 1917 *J. Amer. Chem. Soc.* **39**, 1013.
Johnson and Moelwyn-Hughes 1940 *Proc. Roy. Soc* A, **175**, 118.
Lowry 1899 *J. Chem. Soc.* **75**, 212.
Moelwyn-Hughes 1928 *Trans. Faraday Soc.* **24**, 309.
— 1934 *Z. phys. Chem.* B, **26**, 281.
— 1938*a* *Trans. Faraday Soc.* **34**, 91.
— 1938*b* *Proc. Roy. Soc.* A, **164**, 295.
Richards, Lowry and Faulkner 1927 *J. Chem Soc.* p. 1733.
Riiber 1923 *Ber. dtsch. chem. Ges.* **56**, 2185.
Smith and Smith 1937 *J. Chem. Soc.* p. 1413.

Inner excited states of the proton and neutron

By W. Heitler and S. T. Ma
H. H. Wills Physical Laboratory, University of Bristol

(*Communicated by N. F. Mott, F.R.S.—Received* 14 *May* 1940)

To avoid the difficulties of applying the present meson theory to the interaction of fast mesons with nuclear particles it is proposed as a hypothesis that the charge and spin of an electron can assume higher quantum states The theory is developed and the possibility of observing the new particles discussed.

1. Introduction

The fact that the proton and neutron have anomalous magnetic moments of $+2{\cdot}78$ and $-1{\cdot}93$ Bohr nuclear magnetons does not make it very likely that these particles can be described by a simple one-particle wave equation. Indeed, in order to account for the right magnitude of the magnetic moment, it would be necessary to introduce extra terms in Dirac's equation with odd numerical factors corresponding to an 'intrinsic' moment of the above magnitude. This intrinsic moment is not a 'natural' relativistic effect as it is in the case for the electron

In the meson theory it was possible to link up the existence of the anomalous magnetic moments with other effects such as the nuclear forces. According to this theory a proton or neutron produces a meson field similar to the electromagnetic field produced by charged particles. The anomalous magnetic moments are accounted for by the virtual emission of a meson into states with finite angular momentum. Even in this theory it was necessary to attribute to a proton two sorts of field-producing quantities, namely a 'charge' g producing the meson analogue of an electric field and an 'intrinsic magnetic moment' f producing the meson analogue of a magnetic field. The introduction of an intrinsic moment of some sort seems thus to be unavoidable.

Judging from the odd values which the magnetic moments have, it seems likely that the basic mechanism of these intrinsic moments is of a rather complicated nature. It is perhaps not unplausible to assume that a heavy particle has certain *inner degrees of freedom* and that its spin is only the lowest of a whole series of quantum states arising from one of these degrees of freedom. There are, in fact, a number of arguments in favour of such a hypothesis:

The meson theory in its present form exhibits a number of serious difficulties if applied to the interaction of fast mesons with a nuclear particle. This interaction increases rapidly with increasing energy, and hence all quantities involving fast mesons diverge or are far too big to fit the experiments. The anomalous magnetic moment of the proton, for instance, diverges strongly. The cross-section for the scattering of fast mesons by nuclei is found to increase rapidly with energy which is contrary to the experiments, and even at small energies the cross-section is bigger by an order of magnitude than experiments permit.

A closer analysis (Heitler 1939; cf. also Heisenberg 1939) shows now that the reasons for these discrepancies lie in the fact that in the meson theory a nuclear particle has two degrees of freedom with which a meson has a very strong interaction and which can vibrate with a comparatively small inertia. These degrees of freedom are the spin and the charge. The non-relativistic interaction between a meson field and a nuclear particle with spin $\boldsymbol{\sigma}$ is*

$$H = \Pi\left(\frac{g}{\lambda}\operatorname{div}\boldsymbol{\psi} + \frac{2f}{\lambda}(\boldsymbol{\sigma}\operatorname{curl}\boldsymbol{\phi})\right) + \text{compl. conj.} \qquad (1)$$

where $f/\lambda = f\hbar/\mu c$ is the intrinsic magnetic moment (in the sense of the meson theory). If we compare this formula with the magnetic moment of a proton $e\hbar/2Mc$, we see that the inert mass responsible for any change of the spin direction under the influence of a meson field is the small meson mass μ and not, as one would expect, the proton mass.

Similar considerations can be applied to the charge. The possibility of a change of charge of the heavy particle by emission or absorption of a meson also represents a new degree of freedom. This degree of freedom gives rise to two quantum states for the charge with eigenvalues $+1$ (proton) and 0 (neutron). No attempt has been made so far at a classical description of the charge, and it is not quite clear which quantity shall be defined as the 'inert mass' of the charge, but it is clear that the exchange of charge takes place with a certain inertia.

The experiments show now that all effects arising from the meson-proton (or neutron) interaction are, if not diverging, too big by an order of magnitude. This can only be remedied if we attribute a *larger inertia* to these two degrees of freedom.

In quantum theory this can be done by assuming that further quantum states exist arising from these two degrees of freedom, or, in other words, that states with higher charge and with higher spin exist. It has already been shown in a preliminary note by one of us (Heitler 1940) that by the introduction of those states many of the difficulties of the present meson theory disappear. We shall therefore assume that a 'proton' can have higher charge states *with charges* $+2e$, $+3e$, …, *and also with negative charges* $-e$, $-2e$, …. Similarly, we assume that ***higher spin states*** exist *with spins* $\frac{3}{2}$, $\frac{5}{2}$, …, *etc., for each charge.* Since particles with the above properties have not been observed so far, we shall have to assume that they have a somewhat higher rest energy than the proton. It will be seen in this paper that the first excited states (spin $\frac{3}{2}$ and charge $-e$ and $+2e$) have excitation energies of the order of magnitude of 20 MeV. The fact that these excitation energies are still small compared with the rest energy of the proton Mc^2 makes it possible to treat the higher states in a *non-relativistic* approximation.

---

* Notation as in Fröhlich, Heitler and Kemmer (1938). The factor 2 in the second term of (1) has been added because in this paper we shall denote by $\boldsymbol{\sigma}$ the spin vector with the eigenvalues $\pm\frac{1}{2}$, whereas in the paper by Fröhlich *et al.* $\boldsymbol{\sigma}$ was twice this value.

It is probable, however, that the relativistic generalization will not meet with difficulties of any principal nature, at least as far as the higher spin states are concerned. Dirac (1936) and Fierz and Pauli (1939) have shown that relativistic wave equations for particles with arbitrary spin are possible in principle, although they become increasingly complicated for high spins.*

In this paper we shall confine ourselves to the non-relativistic case (non-relativistic with respect to the heavy particle not with respect to the meson).

The introduction of the higher states also affects the nuclear potentials at small distances. The most important result is that in this theory there are also strong attractive proton-proton forces even without the introduction of the so-called neutretto (neutral meson), and it is possible that the proton-proton scattering data can be explained without this particle which has not been observed in cosmic radiation so far. We shall therefore develop our theory introducing charged mesons only. Neutrettos could be included quite simply if this should be required by the experiments.

On the whole it will be seen that the new theory leads to results consistent with the experiments for all effects depending on the interaction of mesons with nuclear particles, and that no more serious difficulties occur than, for instance, in the theory of radiation.

---

* In view of the results obtained in this paper, it would seem perhaps more natural to try and find a field theory describing particles with all spins simultaneously rather than to build up field theories describing particles with one given spin only.

2. The spin rotator

In order to introduce states with higher spin we use the simplest possible model. We assume that the spin of a proton or neutron is described by a *rotator* which differs from an ordinary rotator only in that the quantum numbers j are half-integer instead of integer. Let $\boldsymbol{\sigma}$ be the vector of the angular momentum (divided by $\hbar$). The components of $\boldsymbol{\sigma}$ satisfy the commutation relations

$$\sigma_x\sigma_y-\sigma_y\sigma_x=-i\sigma_z, \tag{2}$$

and the eigenvalues of σ^2 are

$$\sigma^2=j(j+1), \quad j=\tfrac{1}{2},\tfrac{3}{2},\ldots.$$

For spin $j=\frac{1}{2}$ the stronger condition $\sigma_x\sigma_y+\sigma_y\sigma_x=0$ is valid, but for higher spins this is no longer the case. In addition to $\boldsymbol{\sigma}$ we introduce a *co-ordinate vector* $\mathbf{s}$, the physical significance of which will be that of the *intrinsic magnetic dipole* (in the sense of the meson theory). The components of $\mathbf{s}$ commute with each other

$$s_xs_y-s_ys_x=0 \tag{3}$$

and satisfy with $\boldsymbol{\sigma}$ the usual commutation relations

$$\sigma_xs_y-s_y\sigma_x=is_z, \quad \sigma_xs_z-s_z\sigma_x=-is_y. \tag{4}$$

By the fact that $\boldsymbol{\sigma}$ and $\mathbf{s}$ are vectors with respect to spacial rotations, the matrix elements of $\boldsymbol{\sigma}$ and $\mathbf{s}$ are determined to a large extent. We use a representation in which σ^2 and σ_z are diagonal matrices and have eigenvalues $j(j+1)$ and m respectively. Then we have (cf. for instance, Born and Jordan 1930, pp. 150–2)

$$(\sigma_x+i\sigma_y)_{jj}^{m+1,m}=(\sigma_x-i\sigma_y)_{jj}^{m,m+1}=\sqrt{\{(j+m+1)(j-m)\}}, \tag{5a}$$

$$\sigma_{zjj}^{m}=m, \tag{5b}$$

$$\mathbf{s}_{jj}=a_j\,\boldsymbol{\sigma}_{jj}, \tag{5c}$$

$$s_{zj,j-1}^{mm}=-b_j\sqrt{(j^2-m^2)}, \tag{5d}$$

$$(s_x+is_y)_{j,j-1}^{m+1,m}=b_j\sqrt{\{(j+m+1)(j+m)\}}, \tag{5e}$$

$$(s_x+is_y)_{j-1,j}^{m+1,m}=-b_j^\dagger\sqrt{\{(j-m-1)(j-m)\}}. \tag{5f}$$

a_j and b_j are independent of m. The commutability (3) of the components of s leads to two further relations between a_j, b_j, viz.

$$a_j^2+(2j-1)\,b_jb_j^\dagger-(2j+3)\,b_{j+1}^\dagger b_{j+1}=0, \tag{6a}$$

$$(j+1)\,a_jb_j-(j-1)\,b_ja_{j-1}=0. \tag{6b}$$

The formulae (5) and (6) are valid for any system satisfying the commutation relations (2)–(4). In general the a_j and b_j will be matrices. The special features of a simple rotator are expressed in the fact that to each value of j there is only one quantum state. The a_j, b_j are then pure numbers which commute with each other and which we can assume to be real, $b_j = b_j^\dagger$. The recurrence formulae (6) can then readily be solved. The result depends, however, essentially on whether j is integer or half-integer. For half-integer j we find

$$a_j = \frac{3}{4}\frac{1}{j(j+1)}, \tag{7a}$$

$$b_j = \frac{3}{4j}. \tag{7b}$$

a_j is normalized so that $a_{\frac{1}{2}} = 1.$

$\mathbf{s}_{\frac{1}{2}\frac{1}{2}}$ is then identical with $\boldsymbol{\sigma}_{\frac{1}{2}\frac{1}{2}}$.

According to (5c) the jj matrix elements of s are, apart from the constant a_j, identical with the matrix elements of σ. This is in contrast to the well-known facts for integer spin. In the latter case the jj elements of the dipole moment all vanish, as can also be seen immediately from the recurrence formulae (6). For large j the difference disappears as a_j decreases rapidly with increasing j.

$\mathbf{s}$ could so far be a vector or a pseudovector. $\boldsymbol{\sigma}$ certainly is a pseudovector, i.e. an antisymmetrical tensor as can be seen from (2). The behaviour of $\mathbf{s}$ with respect to a reflexion at the origin depends on what we assume about the behaviour of a_j and b_j. We have assumed that for $j = \frac{1}{2}$ $\mathbf{s}$ and $\boldsymbol{\sigma}$ are identical. This means that $\mathbf{s}$ also is a pseudovector. The a_j, b_j are then invariant against reflexions at the origin and $\mathbf{s}$ has the transformation properties of a *magnetic dipole moment*. The square of $\mathbf{s}$ has only diagonal elements and is proportional to the unit matrix. We easily find

$$s^2 = \tfrac{9}{4}. \tag{8}$$

The fact that $\mathbf{s}$ is identical with $\boldsymbol{\sigma}$ for $j = \frac{1}{2}$ enables us now to generalize the interaction between a meson and a heavy particle in such a way that transitions to higher spin states are possible. We replace in (1) the spin vector $\boldsymbol{\sigma}$ by the magnetic dipole moment $\mathbf{s}$. Furthermore, we have to add a term containing the excitation energy Δ_j for the higher spin states. Thus we assume that the Hamiltonian is

$$H = \Delta_j + \Pi\left(\frac{g}{\lambda}\operatorname{div}\boldsymbol{\psi} + \frac{2f}{\lambda}(\mathbf{s}\operatorname{curl}\boldsymbol{\phi})\right) + \text{compl. conj.} \tag{9}$$

instead of (1). We normalize Δ_j so that $\Delta_{\frac{1}{2}} = 0$.

As to the dependence of Δ_j on j a plausible suggestion can be made. The rotator model used above suggests a formula of the type

$$\Delta_j = \frac{\hbar^2}{2Mr^2}\{j(j+1)-\tfrac{3}{4}\} \equiv \Delta\{j(j+1)-\tfrac{3}{4}\}.$$

The term $\frac{3}{4}$ has been added in order that $\Delta_{\frac{1}{2}} = 0$. Inserting for M the proton mass and for r the electronic radius, Δ is of the order of magnitude of 5 MeV; $\Delta_{\frac{3}{2}}$ would then be 15 MeV We shall see below that the experiments lead to a value for this difference of the order of 20 MeV, which is just the right order of magnitude. For the following applications we shall need only the first excitation energy $\Delta_{\frac{3}{2}}$, and for this we shall use the value determined from experiments.

3. HIGHER CHARGE STATES

In the absence of any classical analogy and any theory of the elementary charge there is no simple model which we could use as a guidance for the introduction of the higher charge states. We shall have to introduce them in a more or less phenomenological way: In (1) Π^* is the operator changing a proton into a neutron and is zero if applied to a neutron, Π is the operator changing a neutron into a proton and is zero if applied to a proton. We now change the definition of Π^* and Π so that: *Π^* is an operator changing a heavy particle with charge ne into a particle with charge $(n-1)e$ and Π is an operator changing a particle with charge ne into a particle with charge $(n+1)e$.* By emission or absorption of mesons a proton can then go over into a particle with charge $2e$ and a neutron into a particle with charge $-e$. Particles with charge ne can, of course, also have all different spins j. In the following we denote particles with positive charge ne and spin j by $\mathrm{P}_j^{(n)}$ and particles with negative charge $-ne$ or charge 0 by $\mathrm{N}_j^{(-n)}$ and $\mathrm{N}_j^{(0)}$ respectively.

Particles with charge ne will have an excitation energy depending on n except if $n = 0$ or 1. From the theoretical point of view nothing can be said about the dependence of the excitation energy on n. In the following sections we shall make use only of the first excited state ($n = +2$ and $n = -1$), and for this we shall determine the order of magnitude of the excitation energy from the experimental facts. The excitation energy for the states $\mathrm{P}_j^{(n)}$ and $\mathrm{N}_j^{(-n)}$ will be denoted by $\Delta_j^{(n)}$ and $\Delta_j^{(-n)}$ with $\Delta_{\frac{1}{2}}^{(1)} = \Delta_{\frac{1}{2}}^{(0)} = 0$.

4. THE ANOMALOUS MAGNETIC MOMENTS OF THE PROTON AND NEUTRON

As a first application of our theory we calculate the anomalous magnetic moment of a proton in the normal state with spin $j = \frac{1}{2}$ and charge $+1e$, using the usual perturbation method. We assume a proton with spin in the $+z$-direction placed in a weak external magnetic field **H** in the same direction and compute the change of the self-energy under the influence of the magnetic field. In the old meson theory a similar calculation (Fröhlich *et al.* 1938, which is referred to below as F. H. K.) leads to an expression for the anomalous magnetic moment μ' which diverges like $\int^{\infty} dk$. The divergence is due to the contribution of very fast mesons which are virtually emitted by the heavy particle. This is no longer the case in our new theory. Whilst in the old theory a proton with spin $+\frac{1}{2}$ could only emit a positive meson and only into states with angular momentum $+1\hbar$, it can, in the new theory, also emit negative mesons, leaving the proton in an excited state with charge $+2e$. Furthermore, positive and negative mesons can be emitted into states with angular momentum $-1\hbar$. In the external magnetic field negative mesons and mesons with angular momentum $-1\hbar$ give, of course, contributions to the magnetic moment with opposite sign. If we would neglect the excitation energies, the resulting magnetic moment would be zero. The finite value of the excitation energies leads to a finite value of the magnetic moment, but the contribution of fast mesons with energies large compared with the excitation energies will cancel out. As a result it will be found that the magnetic moment only diverges logarithmically.

For the calculation we adopt the same method which was applied in F. H. K.; we expand the meson field into spherical waves and take into account that a positive meson with angular momentum $+1\hbar$ about the z-axis has a magnetic energy in the magnetic field of $e\hbar cH/2\epsilon$, where ϵ is the relativistic energy of the meson.* The self energy is given by

$$W = -\sum_n \frac{H_{0n} H_{n0}}{E_n}, \tag{10}$$

where E_n is the energy of the meson in the intermediate state. For H_{0n}, H_{n0} we have to take the matrix elements of (9) for the emission and reabsorption of a meson by the heavy particle respectively. Apart from the fact that $\boldsymbol{\sigma}$ is now replaced by $\mathbf{s}$, the matrix elements of (9) are identical with those of (1) and can be taken directly from F. H. K. The sum (10) has now to be extended over six groups of intermediate states, namely (i) emission of a positive meson into a state with angular momentum about the z-axis $+1\hbar$, leaving the heavy particle in the normal state $\mathrm{N}_{\frac{1}{2}}^{(0)}$ with spin component in the z-direction $m_{\mathrm{P}} = -\frac{1}{2}\hbar$. This was the only term occurring in the old theory; (ii) emission of a negative meson into a state with angular momentum $+1\hbar$, leaving the heavy particle in a state $\mathrm{P}_{\frac{1}{2}}^{(2)}$ with $m_{\mathrm{P}} = -\frac{1}{2}\hbar$, (iii) and (iv) emission of a positive or negative meson into a state with angular momentum about the z-axis $+1\hbar$ but leaving the heavy particle in a state $\mathrm{N}_{\frac{3}{2}}^{(0)}$ or $\mathrm{P}_{\frac{3}{2}}^{(2)}$ with $m_{\mathrm{P}} = -\frac{1}{2}\hbar$; (v) and (vi) emission of a positive or negative meson into a state with angular momentum $-1\hbar$, leaving the heavy particle in a state $\mathrm{N}_{\frac{3}{2}}^{(0)}$ or $\mathrm{P}_{\frac{3}{2}}^{(2)}$ with $m_{\mathrm{P}} = \frac{3}{2}\hbar$. A similar set of intermediate states arises from pair creation. The contribution from these states is the same and results simply in a factor 2.

---

* The method of expansion into spherical waves is simpler and, from the physical point of view, clearer than the usual method of expansion into plane waves treating the influence of the magnetic field as a perturbation in Born's approximation. Both methods lead to exactly the same result. In F. H. K. there was a mistake of a factor 2 which is just compensated by the contribution from pair creation which was not taken into account in that paper. The final formula given in F. H. K. is correct.

The denominators of (10) for these six cases are:

$$\left.\begin{array}{lll}
\text{(i)} & & E_n = \hbar c\sqrt{(k^2+\lambda^2)} - \dfrac{e\hbar cH}{2\epsilon}, \\[2ex]
\text{(ii)} & & E_n = \hbar c\sqrt{(k^2+\lambda^2)} + \dfrac{e\hbar cH}{2\epsilon} + \Delta_{\frac{1}{2}}^{(2)}, \\[2ex]
\text{(iii), (v)} & & E_n = \hbar c\sqrt{(k^2+\lambda^2)} \pm \dfrac{e\hbar cH}{2\epsilon} + \Delta_{\frac{3}{2}}^{(0)},\ \Delta_{\frac{3}{2}}^{(2)}, \\[2ex]
\text{(iv), (vi)} & & E_n = \hbar c\sqrt{(k^2+\lambda^2)} \mp \dfrac{e\hbar cH}{2\epsilon} + \Delta_{\frac{3}{2}}^{(0)},\ \Delta_{\frac{3}{2}}^{(2)}.
\end{array}\right\} \tag{11}$$

Here k is the wave number of the emitted meson. We write $(\mathbf{s}\,\mathrm{curl}\,\boldsymbol{\phi})$ in the form

$$\begin{aligned}(\mathbf{s}\,\mathrm{curl}\,\boldsymbol{\phi}) = s_z\,\mathrm{curl}_z\,\boldsymbol{\phi} + \tfrac{1}{2}(s_x+is_y)\,(\mathrm{curl}_x - i\,\mathrm{curl}_y)\,\boldsymbol{\phi} \\ + \tfrac{1}{2}(s_x-is_y)\,(\mathrm{curl}_x + i\,\mathrm{curl}_y)\,\boldsymbol{\phi}.\end{aligned} \tag{12}$$

In the matrix element H_{0n} of (10) the expression (12) has to be taken at the position of the heavy particle. As explained in F. H. K. (p. 172, equation (57)) the first term is always zero, the second term is different from zero only for a state with angular quantum number $m = -1$, and the third term only for $m = +1$. In the two cases we have

$$\left.\begin{array}{ll}
m = +1 & (\mathbf{s}\,\mathrm{curl}\,\boldsymbol{\phi})_0 = +\tfrac{1}{2}(s_x - is_y)\,ic\dfrac{2\sqrt{2}}{\sqrt{3}}\,k^2q, \\[2ex]
m = -1 & (\mathbf{s}\,\mathrm{curl}\,\boldsymbol{\phi})_0 = -\tfrac{1}{2}(s_x + is_y)\,ic\dfrac{2\sqrt{2}}{\sqrt{3}}\,k^2q.
\end{array}\right\} \tag{13}$$

In a similar way we find for the conjugate matrix elements

$$\left.\begin{aligned} m&=+1 \quad (\mathbf{s}\,\mathrm{curl}\,\boldsymbol{\phi})_0^* = -\tfrac{1}{2}(s_x+is_y)\,ic\frac{2\sqrt{2}}{\sqrt{3}}k^2q^*,\\ m&=-1 \quad (\mathbf{s}\,\mathrm{curl}\,\boldsymbol{\phi})_0^* = +\tfrac{1}{2}(s_x-is_y)\,ic\frac{2\sqrt{2}}{\sqrt{3}}k^2q^*.\end{aligned}\right\} \tag{14}$$

It has to be noted that $m\hbar$ is the angular momentum about the z-axis only for a positive meson, for a negative meson this quantity is $-m\hbar$ because wave number and momentum have then opposite signs.

q has matrix elements for the emission of a negative meson or the absorption of a positive one. q^* has matrix elements for the emission of a positive meson and the absorption of a negative one. q and q^* are given by equations (25)–(29) in F. H. K.

We divide the product of the matrix elements $H_{0n}H_{n0}$ for the six intermediate states by the corresponding denominators (11) and replace the sum Σ by an integral $\int dk/\pi$. Expanding W into a power series with respect to H, we obtain for the term proportional to H

$$W = \frac{4}{3\pi}\frac{f^2}{\lambda^2}\hbar^3c^3eH\{\quad\}, \tag{15}$$

where

$$\begin{aligned}\{\quad\} = {} & (s_x+is_y)_{j,j}^{m,m-1}(s_x-is_y)_{j,j}^{m-1,m}\left[\frac{1}{\epsilon^2}-\frac{1}{(\epsilon+\Delta_1^{(2)})^2}\right]\\ & +(s_x+is_y)_{j,j+1}^{m,m-1}(s_x-is_y)_{j+1,j}^{m-1,m}\left[\frac{1}{(\epsilon+\Delta_1^{(0)})^2}-\frac{1}{(\epsilon+\Delta_1^{(2)})^2}\right]\\ & -(s_x-is_y)_{j,j+1}^{m,m+1}(s_x+is_y)_{j+1,j}^{m+1,m}\left[\frac{1}{(\epsilon+\Delta_1^{(0)})^2}-\frac{1}{(\epsilon+\Delta_1^{(2)})^2}\right],\end{aligned} \tag{16}$$

with $j = m = \frac{1}{2}$. If we would neglect the excitation energies Δ, (16) would vanish and that for two reasons: first, the contributions of positive and negative mesons would cancel each other, and, secondly, even if no higher charge states and only positive mesons would exist, (16) would be proportional to

$$[(s_x+is_y)(s_x-is_y)]_{jj}^{mm}-[(s_x-is_y)(s_x+is_y)]_{jj}^{mm},$$

which vanishes because the components of s commute. It will be seen below that the Δ's are small compared with the rest energy of the meson. We can thus expand (16) for $\Delta \ll \epsilon$. Inserting the matrix elements for the s from (5), we have

$$\{\quad\} = \frac{2}{\epsilon^3}(\Delta_1^{(2)} + \Delta_1^{(0)} - \Delta_1^{(2)}) \equiv \frac{2}{\epsilon^3}\Delta. \tag{17}$$

Dividing (15) by H and by the Bohr nuclear magneton $\mu_0 = e\hbar/2Mc$, we find for the anomalous magnetic moment μ'

$$\frac{\mu'}{\mu_0} = \frac{16}{3\pi}\frac{M}{\mu}\frac{f^2}{\hbar c}\frac{\Delta}{\mu c^2}\int_0^\infty \frac{k^4\,dk}{(k^2+\lambda^2)^{\frac{5}{2}}},$$

where μ is the meson mass. The integral diverges logarithmically. We integrate to an upper limit E_m for the energy $\hbar c\sqrt{(k^2+\lambda^2)}$ large compared with μc^2 and obtain

$$\frac{\mu'}{\mu_0} = \frac{16}{3\pi}\frac{M}{\mu}\frac{f^2}{\hbar c}\frac{\Delta}{\mu c^2}\left(\log\frac{2E_m}{\mu c^2} - \frac{4}{3}\right). \tag{18}$$

The divergence of the magnetic moment is of a lesser degree than in the old theory, but it has not disappeared entirely yet. This is, at the present state of our theory, hardly to be expected. Our theory is as far as the heavy particle is concerned non-relativistic It is well known that in Dirac's hole theory the divergence of the self energies is less serious than in a non-relativistic theory. There are therefore reasons to believe that the magnetic moment will converge if the relativistic features of the heavy particle are taken into account. For the limit of validity of our theory we can therefore assume the rest energy of the proton $E_m = Mc^2$. From (18) and from the known value of $\mu'/\mu_0 \sim 2$ we can then determine the product $f^2\Delta$ and find for the dimensionless quantity

$$\frac{f^2}{\hbar c}\frac{\Delta}{\mu c^2} \propto \frac{1}{20}. \tag{19}$$

In §§ 5 and 6 we shall find further relations for the constants f and the Δ's enabling us to determine them separately. Presumably, the three Δ's in (17) are all of the same order of magnitude and for the present we may assume that they are all equal. It will be seen that the nuclear potentials and the cross-section for scattering are of the right order of magnitude for $g^2/\hbar c = f^2/\hbar c = 0{\cdot}08$ and $\Delta \sim 20$ MeV. This would give a value for

$$f^2\Delta/\hbar c\mu c^2 \propto 1/50$$

which is the reasonable agreement with (19), considering the uncertainty as to the exact determination of the constants.

5. THE SCATTERING OF MESONS

Owing to the interaction between the meson field and a nuclear particle, a meson can be scattered by a free proton or neutron. The cross-section for this scattering has been calculated according to the meson theory by Heitler (1938) and Bhabha (1938).*

The cross-section for the scattering of a longitudinal meson with momentum p/c and energy $\epsilon < Mc^2$ by a free proton or neutron at rest was found to be

$$\Phi = 4\pi\left(\frac{g^2}{\mu c^2}\right)^2 \frac{p^4}{\epsilon^2(\mu c^2)^2}. \tag{20}$$

This formula has two very interesting features First, the mass of the heavy particle M does not occur in (20) at all, whilst, from the analogy of mesons with light quanta, one would expect a formulae of the type of the Thomson formula, viz. $\Phi = \frac{8\pi}{3}(g^2/Mc^2)^2$. (Cf. Nordheim and Teller 1938). Instead of M (20) contains the meson mass μ in the denominator. Secondly, (20) increases rapidly with energy for $\epsilon > \mu c^2$. If this were true, multiple processes would occur at high energies. A single meson could produce a shower of mesons in a single collision with a heavy particle The reason for the different behaviour of mesons and light quanta is the following: The scattering of a light quantum by a free charged particle takes place by excitation of a certain degree of freedom of the particle, and this degree of freedom is the free motion. The cross-section is inversely proportional to the square of the inert mass of the scattering particle and is independent of the energy as long as the latter is smaller than the rest energy of the scattering particle. The scattering of mesons cannot be due to this same degree of freedom, since the mass M of the scattering particle does not occur in (20). In fact the scattering as given by the cross-section (20) is due to another degree of freedom, namely, the charge. This can be seen from the fact that neutral mesons would be scattered according to a much smaller cross-section. This has, for instance, been shown by Bhabha (1939), who has developed a classical theory for neutral mesons. We shall arrive at the same result below. In addition to (20) there is, of course, also a contribution to the scattering from free motion, but this contribution is of the order of the Thomson cross-section $(g^2/Mc^2)^2$, which is small compared with (20). For transverse mesons a second degree of freedom, the spin, contributes to the scattering with a cross-section of the same order of magnitude as (20).

---

* In the scalar meson theory a similar calculation had already been made by Yukawa and Sakata (1937).

(20) is in violent disagreement with the experimental facts. Recently, J. G. Wilson (1940) has estimated the cross-section for the scattering of mesons in the energy region between a few times 10^8 and 10^9 eV. For low energies the cross-section was found to be less than 10^{-27} cm.^{2} and for 10^9 eV the cross-section is of the order of 10^{-28} cm.^{2}. There is no indication for an increase of Φ with energy nor is there any indication for multiple processes to occur in this energy region. Even for an energy as low as $\epsilon = 2\mu c^2$, (20) would be of the order of 10^{-26} cm.^{2}, which is certainly out of the question.

In our new theory the inertia of the two degrees of freedom in question—charge and spin—is very much larger and consequently the cross-section for scattering will be very much smaller than (20) and in agreement with the experiments. Below we calculate the contribution to the scattering from these two degrees of freedom.

It will be seen that it is not much larger than the contribution from the free motion and a complete expression for the cross-section would have to include the latter as well. This, however, is only partially possible in our new theory, at least in its present—very preliminary—state. The interaction (8) is only valid for a heavy particle at rest, but for the calculation of the scattering due to free motion the interaction with a moving particle is required, at least up to terms proportional to v/c. In the old meson theory these terms can easily be derived from the relativistic generalization of (1), but we do not possess the relativistic generalization for the spin dependent part of (8) in the new theory. Although there should be no principal difficulty in finding the interaction of a slowly moving dipole (according to our model) with a meson field up to terms $\propto v/c$, we shall not be concerned with this question in this paper. No difficulty of this sort occurs for the scattering of mesons due to the spin independent g-type of interaction. For this case we shall give the contribution to the cross-section from the free motion below.

The following calculations are correct to the extent to which a proton or neutron can be considered as infinitely heavy.

We consider the scattering of a positive primary meson with momentum $\mathbf{p}_0/c$ by a free neutron at rest. The meson will be scattered into a state with momentum $\mathbf{p}/c$. We neglect the recoil and have therefore $|p_0| = |p|$. We calculate the cross-section, using Born's approximation. The scattering can then be considered as a two-stage process with two sets of intermediate states: (i) The primary meson is first absorbed, transforming the neutron into a proton and, subsequently, the proton emits the secondary meson, returning itself into the normal neutron state. (ii) The secondary meson is first emitted by the neutron, which is transformed into an excited $N^{(-1)}$

particle. Subsequently, the primary meson is absorbed and the heavy particle returns to the normal $N^{(0)}$ state. If both the primary and secondary mesons are transverse, transitions to higher spin states are also possible and the heavy particle in the intermediate may be in a $P_{\frac{1}{2}}^{(1)}$ or $N_{\frac{1}{2}}^{(-1)}$ state respectively. In this case we have altogether four intermediate states.

We first consider the case where the primary and secondary mesons are longitudinal. The matrix element responsible for the scattering is of the type

$$V = \frac{H_{A\mathrm{I}} H_{\mathrm{I}F}}{E_A - E_{\mathrm{I}}} + \frac{H_{A\mathrm{II}} H_{\mathrm{II}F}}{E_A - E_{\mathrm{II}}},$$

where A, F, I, II are the initial, final, first and second intermediate states respectively. The energy differences are obviously

$$E_A - E_{\mathrm{I}} = \epsilon_0, \quad E_A - E_{\mathrm{II}} = -(\epsilon + \Delta_{\frac{1}{2}}^{(-1)}), \tag{21}$$

where ϵ_0 and ϵ are the energies of the primary and secondary mesons. Inserting for $H_{A\mathrm{I}}$, etc. the matrix elements for the absorption and emission of a longitudinal meson (F. H. K. equation (32)), we have, putting $p_0 = p$ and $\epsilon_0 = \epsilon$,

$$V = \frac{2\pi g^2}{\epsilon\, \lambda^2} p^2 \left(\frac{1}{\epsilon} - \frac{1}{\epsilon + \Delta_{\frac{1}{2}}^{(-1)}} \right). \tag{22}$$

It is interesting to compare this result with the one obtained from the old meson theory. In the absence of excited charge states only the first term $1/\epsilon$ of (22) is there because of the conservation of charge. The cross-section deduced from this matrix element would then be equal to (20). In the new theory V would vanish if we would neglect the excitation energy $\Delta_{\frac{1}{2}}^{(-1)}$. Thus, there would be no contribution to the scattering from the charge degree of freedom. The same is obviously true for the scattering of neutral mesons. Then also both intermediate states occur (even according to the old meson theory) and the only contribution to the scattering is that from free motion.

Since $\Delta \ll \mu c^2 < \epsilon$, we can expand (22) and obtain for V

$$V = 2\pi \frac{g^2}{\lambda^2} \frac{\Delta_{\frac{1}{2}}^{(-1)}}{\epsilon^3},$$

and for the corresponding cross-section

$$\Phi_{\text{long.} \to \text{long}} = 4\pi \left(\frac{g^2}{\mu c^2} \right)^2 \left(\frac{\Delta_{\frac{1}{2}}^{(-1)}}{\mu c^2} \right)^2 \frac{p^4}{\epsilon^4}. \tag{23a}$$

In a similar way we obtain the cross-section for the scattering of a longitudinal meson into a transverse one

$$\Phi_{\text{long.}\to\text{trans.}} = 8\pi\left(\frac{gf}{\mu c^2}\right)^2\left(\frac{\Delta_{\frac12}^{(-1)}}{\mu c^2}\right)^2\frac{p^4}{\epsilon^4}. \tag{23b}$$

Before discussing this result we consider the scattering of a transverse meson. In this case also the higher spin states have to be taken into account. The matrix element is then

$$V = \frac{2\pi}{\epsilon}\left(\frac{2f}{\lambda}\right)^2 p^2\left[\frac{(\mathbf{s}[\mathbf{k}_0\mathbf{j}_0])\,(\mathbf{s}[\mathbf{kj}])}{\epsilon} + \frac{(\mathbf{s}[\mathbf{k}_0\mathbf{j}_0])\,(\mathbf{s}[\mathbf{kj}])}{\epsilon-\Delta_{\frac12}^{(1)}}\right.$$
$$\left. - \frac{(\mathbf{s}[\mathbf{kj}])\,(\mathbf{s}[\mathbf{k}_0\mathbf{j}_0])}{\epsilon+\Delta_{\frac12}^{(-1)}} - \frac{(\mathbf{s}[\mathbf{kj}])\,(\mathbf{s}[\mathbf{k}_0\mathbf{j}_0])}{\epsilon+\Delta_{\frac12}^{(-1)}}\right]. \tag{24}$$

Here $\mathbf{k}_0$ and $\mathbf{k}$ are unit vectors in the direction of $\mathbf{p}_0$ and $\mathbf{p}$, $\mathbf{j}_0$ and $\mathbf{j}$ are unit vectors in the direction of polarization of the primary and secondary mesons respectively. It is easy to see that (24) vanishes if we neglect the Δ's. If we denote by s_0 and s the components of the vector $\mathbf{s}$ in the directions of the unit vectors $[\mathbf{k}_0\mathbf{j}_0]$ and $[\mathbf{kj}]$ respectively, the square bracket of (24) is*

$$[\quad] = \frac{s_{0\frac12\frac12}s_{\frac12\frac12}}{\epsilon} + \frac{s_{0\frac12\frac12}s_{\frac12\frac12}}{\epsilon-\Delta_{\frac12}^{(1)}} - \frac{s_{\frac12\frac12}s_{0\frac12\frac12}}{\epsilon+\Delta_{\frac12}^{(-1)}} - \frac{s_{\frac12\frac12}s_{0\frac12\frac12}}{\epsilon+\Delta_{\frac12}^{(-1)}}. \tag{25}$$

If we now neglect the Δ's, (25) becomes simply

$$[\quad] = \frac{1}{\epsilon}\{(s_0s)_{\frac12\frac12} - (ss_0)_{\frac12\frac12}\}$$

by the laws of matrix multiplication. This expression vanishes because any two components of $\mathbf{s}$ commute For this result both the higher spin and higher charge states are necessary. For finite Δ we expand again as above, assuming $\Delta \ll \epsilon$. Furthermore, we make use of the following formulae:

$$s_{0\frac12\frac12}s_{\frac12\frac12} - s_{\frac12\frac12}s_{0\frac12\frac12} = +i(\boldsymbol{\sigma}[[\mathbf{k}_0\mathbf{j}_0]\,[\mathbf{kj}]]),$$

$$s_{0\frac12\frac12}s_{\frac12\frac12} - s_{\frac12\frac12}s_{0\frac12\frac12} = -i(\boldsymbol{\sigma}[[\mathbf{k}_0\mathbf{j}_0]\,[\mathbf{kj}]]),$$

$$s_{0\frac12\frac12}s_{\frac12\frac12} + s_{\frac12\frac12}s_{0\frac12\frac12} = \tfrac12([\mathbf{k}_0\mathbf{j}_0]\,[\mathbf{kj}]),$$

$$s_{0\frac12\frac12}s_{\frac12\frac12} + s_{\frac12\frac12}s_{0\frac12\frac12} = ([\mathbf{k}_0\mathbf{j}_0]\,[\mathbf{kj}]).$$

---

* In the following $s_{0\frac12\frac12}s_{\frac12\frac12}$, etc. means: $\sum_{m'} s_{0\frac12\frac12}^{m\,m'}s_{\frac12\frac12}^{m'm}$, m being the spin component in the z-direction.

The matrix element (24) then becomes

$$V = \frac{2\pi}{\epsilon^3}\frac{f^2}{\lambda^2}p^2\{(\Delta_1^{(-1)} + 2\Delta_1^{(+1)} + 2\Delta_1^{(-1)})([\mathbf{k}_0\mathbf{j}_0][\mathbf{kj}]) + 2i(\Delta_1^{(-1)} + \Delta_1^{(+1)} - \Delta_1^{(-1)}) \times (\boldsymbol{\sigma}[[\mathbf{k}_0\mathbf{j}_0][\mathbf{kj}]])\}. \quad (26)$$

From (26) we easily find the cross-section. We take the sum over the spin directions of the heavy particle and over the two directions of polarization of the meson in the final state and the average over these directions in the initial state. The cross-section is then

$$\Phi_{\text{trans}\to\text{trans}} = \frac{8\pi}{3}\left(\frac{f^2}{\mu c^2}\right)^2\frac{p^4}{\epsilon^4}\{(\Delta_1^{(-1)} + 2\Delta_1^{(+1)} + 2\Delta_1^{(-1)})^2 + 2(\Delta_1^{(-1)} + \Delta_1^{(+1)} - \Delta_1^{(-1)})^2\}. \quad (23c)$$

The cross-sections (23a, b, c) are all of the same order of magnitude. The most important features are. (i) The cross-section for scattering does no longer increase with energy but tends to a finite value for large ϵ. (ii) At an energy of the order μc^2 the cross-section is smaller by a factor of the order $(\Delta/\mu c^2)^2$ than (20). Using the numerical values for Δ, g, f used in §§ 4 and 6, the cross-section is of the order of 10^{-27} cm.^{2} for $\epsilon = 2\mu c^2$ and is thus, within reasonable limits, in agreement with the facts. For the comparison with the experiments it has further to be remarked that, by relativistic reasons, the cross-section is likely to become smaller for $\epsilon \sim Mc^2$ in a similar way as the cross-section for the scattering of light by an electron. It may well be that this effect is already appreciable at 10^9 eV.

In addition to (23) there is a contribution to the scattering from the free motion. As mentioned above, we are only in the position to compute this contribution for that part of the interaction which is independent of the spin, i.e. the g-interaction. We give the results without calculation·

$$\Phi_{\text{long}\to\text{long}} = \frac{4\pi}{3}\left(\frac{g^2}{Mc^2}\right)^2\left(\frac{\mu c^2}{\epsilon}\right)^4, \quad (27a)$$

$$\Phi_{\text{trans.}\to\text{trans}} = \frac{8\pi}{3}\left(\frac{g^2}{Mc^2}\right)^2, \quad (27b)$$

as was to be expected from the analogy of light quanta and mesons. The same results have already been obtained by Bhabha (1939) from his classical theory, and by Heitler (1939) for neutral mesons.

Both contributions are even smaller than the contributions from charge and spin. There is no increase with energy. In all probability the contributions from the spin dependent part of the interaction will be, in a correct generalization of our theory, of a similar type as (27).

Since the cross-section for scattering no longer increases with energy, *no multiple processes will occur*. This is also in agreement with the experimental facts. According to Lovell (1939) multiple showers, if they exist at all, are extremely rare and do certainly not play an important role in cosmic radiation. We cannot, of course, exclude that at some very high energy those multiple showers might occur, but within the scope of our theory, i e. up to 1000 MeV, they must be expected to be very rare.

6. THE NUCLEAR FORCES

The existence of the higher excited states also influences the potentials between nuclear particles. This influence can be described in a similar way as the interaction of two atoms is described in the theory of molecules. From each pair of charge and spin states of the two particles a potential curve starts which, at large distances, is not affected by the existence of the other states. From the ground states, for instance, we have the same potential curves as obtained from the meson theory in its old form. As soon, however, as the potential becomes of the same order of magnitude as the excitation energy, viz. 20 MeV, the potential curves starting from different excited states perturb each other. As is well known from molecular theories, this perturbation results in a mutual repulsion of potential curves of the same 'race' (i.e the same symmetries). Thus the lowest potential curve of each race will be further deepened. The smaller the distance is the stronger will be the influence of the higher excited states. *The proton-neutron potential will therefore be deeper and more narrow than according to the* $e^{-\lambda r}/r$ *law* derived from the old meson theory. The most important result, however, is that there is also a *strong attractive proton-neutron interaction* due to the existence of the higher charge states. Such a proton-proton force, as required by the scattering experiments, could, in the old meson theory, only be obtained by introducing a neutral meson (neutretto). There seems now to be a possibility of explaining the proton-proton scattering data without such a neutral particle. This would be an advantage inasmuch as the neutretto has not been observed in cosmic radiation so far, and recent experiments by Lovell (1939) show that their intensity is certainly less than 10 % of the intensity of charged mesons. Although this does not necessarily mean that neutrettos could not exist, it would seem unlikely that their intensity in cosmic radiation should be so small.*

The nuclear potentials can be calculated by the method of secular perturbation which is also used in the theory of molecules. The potential between two nuclear particles at a distance r is given by the formula derived from the meson theory

$$V = \frac{e^{-\lambda r}}{r}(P_{12}+P_{21}) \times \left\{g^2+4f^2\left[(\mathbf{s}_1\mathbf{s}_2)\left(1+\frac{1}{\lambda r}+\frac{1}{\lambda^2 r^2}\right)-\frac{(\mathbf{s}_1\mathbf{r})(\mathbf{s}_2\mathbf{r})}{r^2}\left(1+\frac{3}{\lambda r}+\frac{3}{\lambda^2 r^2}\right)\right]\right\}, \quad (28)$$

which differs from the potential of the old meson theory only in that the significance of P_{12} and $\mathbf{s}_1$, $\mathbf{s}_2$ are different. P_{12} is now the operator by which the charge of the particle 1 decreases by one unit and the charge of the particle 2 increases by one unit $\mathbf{s}_1$, $\mathbf{s}_2$ are no longer the spins of the two particles but are the dipole moments introduced in § 1. (28) has a diagonal

* [*Added in proof.* Recently, it has been suggested by Bethe (1940) that the nuclear forces should be due to the exchange of neutral mesons only. The argument put forward by Bethe is based on the fact that in such a neutral theory a quantitative treatment of the deuteron problem is possible, and that this is not the case in an alternative 'symmetrical' theory which differs from Kemmer's form of the meson theory (Kemmer 1939) in that one of the constants (g) is put equal to zero. Especially, it has turned out that in the 'symmetrical' theory the sign of the quadripole moment of the deuteron is the wrong one. We do not think that such a neutral theory can correspond to reality. The most convincing point of the meson theory—perhaps the only one which gives it preference to any of the older nuclear theories—is Yukawa's relation between the range of nuclear forces and the mass of the cosmic ray particles. This connexion is lost in a neutral theory and the same is true for the connexion between the β-decay of cosmic ray mesons and nuclei and for the qualitative explanation of the anomalous magnetic moments. These difficulties have, of course, also been fully realized by Bethe. We believe, however, that in a preliminary theory like the meson theory which can only claim to have the value of some sort of correspondence principle more weight should be put on a qualitative connected account of many different phenomena than on a quantitative treatment of a single effect. This is all the more the case if such a quantitative treatment is only made possible by a suitably chosen 'cutting-off' procedure which is always to some extent arbitrary We do not think either that the negative part of the above argument is conclusive Bethe's 'symmetrical theory' is a very specialized form of the meson theory in which the constants are given special values. It has been shown by Møller and Rosenfeld (1938) for their version of the meson theory and by one of us (Heitler 1939) for a pure vector theory that any sign of the quadripole moment can be obtained if full use is made of the constants available. The fact that a certain specialized form of the theory gives quantitatively wrong results for the deuteron cannot, in our opinion, be used as part of an argument in favour of another theory which can be made to give right results for the deuteron but disagrees with other experimental facts in points of a far more principal nature.]

matrix element for each pair of charge and spin states and has also matrix elements corresponding to transitions between different charge and spin states. In order to simplify our calculations we confine ourselves to S-states. This is a good approximation to reality, although it is known the ground state of the deuteron is a mixture of a ^{3}S- and ^{3}D$_1$-state. The admixture of the D-state is, however, very small. This has been shown by Alvarez and Bloch (1940), who have measured the magnetic moment of the neutron directly and have found that the magnetic moments in the deuteron are additive within 1·5 %. A simple consideration shows then that the admixture of the D-state is certainly not greater than 10 %, a figure which is still big enough to account for the quadripole moment of the deuteron.* For the deuteron potential the D-state is probably unimportant (cf. § 7 A).

For S-states (28) becomes

$$V = \frac{e^{-\lambda r}}{r}(P_{12}+P_{21})\{g^2+\tfrac{8}{3}f^2(\mathbf{s}_1\mathbf{s}_2)\}. \tag{29}$$

The actual potential between the two particles is derived from a secular equation in which the matrix elements of (29) enter. The rank of the determinant depends upon the number of excited states which are taken into account and is in principle infinite. The whole character of our theory is, however, that of an approximation from large distances and for this case only few of the excited states are important. In the first approximation we obtain the same result as in the old theory. As a second approximation we take into account *one excited* state for each particle. We consider, for instance, the proton-proton and proton-neutron singlet states. Up to the second approximation the following states occur.

P-P(^{1}S) $\quad \mathrm{P}^{(1)}_{\uparrow}\mathrm{P}^{(1)}_{\downarrow}$; $\quad \mathrm{P}^{(2)}_{\uparrow}\mathrm{N}^{(0)}_{\downarrow}$; $\quad \mathrm{P}^{(1)}_{\uparrow}\mathrm{P}^{(1)}_{\downarrow}$.

P-N(^{1}S) $\quad \mathrm{P}^{(1)}_{\uparrow}\mathrm{N}^{(0)}_{\downarrow}$; $\quad \mathrm{P}^{(2)}_{\uparrow}\mathrm{N}^{(-1)}_{\downarrow}$; $\quad \mathrm{P}^{(1)}_{\uparrow}\mathrm{N}^{(0)}_{\downarrow}$.

For both cases the charge wave functions are symmetrical in the two particles, since, for a singlet, the spin wave function is antisymmetrical and, for a S-state, the spacial wave function is symmetrical. The charge wave functions are therefore of the type $\frac{1}{\sqrt{2}}\{u(1)\,v(2)+u(2)\,v(1)\}$ or $u(1)\,u(2)$ according to whether the two particles have the same or different charges. Hence, the diagonal matrix elements of $P = P_{12}+P_{21}$ are all equal to unity if the charges of the two particles differ by one and nought otherwise. The non-diagonal elements are equal to $\sqrt{2}$ for the transition from the state $\mathrm{P}^{(1)}\mathrm{P}^{(1)}$ into the state $\mathrm{P}^{(2)}\mathrm{N}^{(0)}$ and 1 for the transition from $\mathrm{P}^{(1)}\mathrm{N}^{(0)}$ to $\mathrm{P}^{(2)}\mathrm{N}^{(-1)}$.

---

* This figure means that if the wave function of the deuteron is written in the form $\psi_s+\alpha\psi_D$, α is of the order of 0·1.

The spin wave function is

$$\frac{1}{\sqrt{2}}\{\alpha(1)\beta(2)-\alpha(2)\beta(1)\},$$

if the two particles both have spin $\frac{1}{2}$. If they both have spin $\frac{3}{2}$, the spin function is

$$\tfrac{1}{2}\{a(1)d(2)-b(1)c(2)+c(1)b(2)-d(1)a(2)\}, \tag{30}$$

where a, b, c, d are the spin functions of one particle with spin component in the z-axis of $\frac{3}{2}$, $\frac{1}{2}$, $-\frac{1}{2}$, $-\frac{3}{2}$ respectively. The matrix elements of $\mathbf{s}$ for the transitions $\alpha \rightarrow a$, $\alpha \rightarrow b$, etc. are given in equations ($5a$–f). Hence we obtain for the matrix elements of $(\mathbf{s}_1\mathbf{s}_2)$

Diagonal element $(\frac{1}{2}\,\frac{1}{2}) \rightarrow (\frac{1}{2}\,\frac{1}{2})$: -3,

Diagonal element $(\frac{3}{2}\,\frac{3}{2}) \rightarrow (\frac{3}{2}\,\frac{3}{2})$: $-\frac{3}{5}$,

Non-diagonal element $(\frac{1}{2}\,\frac{1}{2}) \rightarrow (\frac{3}{2}\,\frac{3}{2})$. $-3\sqrt{2}$.

Thus the secular equations are:

For P-N(^{1}S):

$$\begin{vmatrix} G-2F-\epsilon & G-2F & -2\sqrt{2}\,F \\ G-2F & \Delta_{\frac{1}{2}}^{(2)}+\Delta_{\frac{1}{2}}^{(-1)}-\epsilon & -2\sqrt{2}\,F \\ -2\sqrt{2}\,F & -2\sqrt{2}\,F & G-\frac{2}{5}F+\Delta_{\frac{1}{2}}^{(1)}+\Delta_{\frac{1}{2}}^{(0)}-\epsilon \end{vmatrix} = 0, \tag{31a}$$

For P-P(^{1}S).

$$\begin{vmatrix} -\epsilon & \sqrt{2}\,(G-2F) & 0 \\ \sqrt{2}\,(G-2F) & \Delta_{\frac{1}{2}}^{(2)}-\epsilon & -4F \\ 0 & -4F & 2\Delta_{\frac{1}{2}}^{(1)}-\epsilon \end{vmatrix} = 0, \tag{31b}$$

where $G = g^2e^{-\lambda r}/r$ and $F = f^2e^{-\lambda r}/r$

In a similar way we find the ^{3}S proton-neutron potential In the same approximation it is described by a determinant with four rows and columns

P-N(^{3}S):

$$\begin{vmatrix} -G-\frac{2}{3}F-\epsilon & G+\frac{2}{3}F & \frac{8}{3}F & \frac{2}{3}\sqrt{10}\,F \\ G+\frac{2}{3}F & \Delta_{\frac{1}{2}}^{(2)}+\Delta_{\frac{1}{2}}^{(-1)}-\epsilon & -\frac{8}{3}F & -\frac{2}{3}\sqrt{10}\,F \\ \frac{8}{3}F & -\frac{8}{3}F & -G+\frac{4}{3}F+\Delta_{\frac{1}{2}}^{(1)}-\epsilon & -\frac{4}{3}\sqrt{\frac{2}{5}}\,F \\ \frac{2}{3}\sqrt{10}\,F & -\frac{2}{3}\sqrt{10}\,F & -\frac{4}{3}\sqrt{\frac{2}{5}}\,F & -G+\frac{22}{75}F+2\Delta_{\frac{1}{2}}^{(1)}-\epsilon \end{vmatrix} = 0 \tag{31c}$$

corresponding to the four states

$$\mathrm{P}^{(1)}_{\frac{1}{2}}\mathrm{N}^{(0)}_{\frac{1}{2}};\quad \mathrm{P}^{(2)}_{\frac{1}{2}}\mathrm{N}^{(-1)}_{\frac{1}{2}};\quad \mathrm{P}^{(1)}_{\frac{1}{2}}\mathrm{N}^{(0)}_{\frac{1}{2}} \text{ and } \mathrm{P}^{(1)}_{\frac{1}{2}}\mathrm{N}^{(0)}_{\frac{1}{2}};\quad \mathrm{P}^{(1)}_{\frac{1}{2}}\mathrm{N}^{(0)}_{\frac{1}{2}}.^{*}$$

In (31*a*–*c*) a number of constants occur upon which the actual potential ϵ will depend. These constants are g^2, f^2 and the Δ's. A variation of these constants gives rise to a great variety of potential curves which, however, are all of the same type if the constants are varied within reasonable limits. We determine the constants from the following experimental facts: The depth of the ^{3}S P-N potential is ~ 25 MeV and that of the ^{1}S P-N potential about half of this figure. For the depth we take the value of the potential at a distance where $e^{-\lambda r}/\lambda r = 1$. Furthermore, we have the value of the anomalous magnetic moment and a rough estimate of the scattering cross-section. The Δ's are probably all of the same order of magnitude and, for the purpose of this section, we can assume that they are all equal. We have found that the following choice of the constants leads to a reasonable agreement with the above experimental facts.

$$g^2\lambda = f^2\lambda = 0{\cdot}3\Delta, \quad \Delta \sim 20 \text{ MeV}. \tag{32}$$

These figures are, of course, very rough and are only intended to indicate the order of magnitude, especially because no exact meaning can be attached to the "depth" of our potential curves. An exact determination would involve the solution of the Schrödinger equation for our potentials. The figures (32) have already been used in §§ 4 and 5 and have been found to give reasonable results for the magnetic moment and the cross-section for the scattering of mesons.

The potentials can then easily be worked out numerically. In figure 1 we give the lowest potential curve for the three cases. For the sake of a simple representation we have plotted the potentials divided by $e^{-\lambda r}/\lambda r$, viz. the function $f(\lambda r)$ defined by

$$\epsilon = f(\lambda r)\, e^{-\lambda r}/\lambda r. \tag{33}$$

In the old meson theory $f(\lambda r)$ is constant (dotted in figure 1; the choice o the constants (32) would, of course, have to be different in the old theory).

The curves obtained are the second approximation coming from large distances. At very small distances the potential curves are further deepened by the influence of still higher excited states. In figure 1 the part of the curves which would be influenced by the higher approximations is dotted. It is, however, quite easy to find a lower limit for the potentials at very small distances. Such a lower limit is obtained if we neglect the excitation energies altogether. As far as the charge is concerned, the secular problem can then be solved exactly. If ne is the charge of one particle, the charge of the other particle is $-(n-2)e$ for the P-P case and $-(n-1)e$ for the P-N case. The secular problem is then equivalent with the solution of the following set of linear equations

---

* In (31*c*) we have assumed $\Delta^{(1)}_{\frac{1}{2}} = \Delta^{(0)}_{\frac{1}{2}}$.

$$(a_{n+1}+a_{n-1}) = pa_n, \tag{34}$$

p being the eigenvalues of the operator $P_{12}+P_{21}$ occurring in (29). There are certain boundary conditions to be satisfied for $n = 1$ which depend on whether we consider the P-P and P-N case. In both cases the eigenvalues p are the same, viz.

$$p = 2\cos k \quad (0 \leqslant k < 2\pi).$$

Thus we have a lower limit for all potentials if we put

$$P_{12}+P_{21} = -2.$$

This is also the *asymptotic value* to which all potential curves tend in the limit $r \to 0$ It is to be noted that this *asymptotic value is charge independent*, i.e. the same for the P-P and P-N case

A lower limit for $(P_{12}+P_{21})(\mathbf{s}_1\mathbf{s}_2)$ can also easily be found. Since the components of s_1 and s_2 all commute, we can choose a representation in which s_1 and s_2 are both diagonal. The lower limit of $(\mathbf{s}_1\mathbf{s}_2)$ is then, according to (8),

$$(P_{12}+P_{21})(\mathbf{s}_1\mathbf{s}_2) > -2\times\tfrac{9}{4}.$$

Thus a lower limit for all potential curves is

$$\epsilon > -2(g^2+6f^2)\frac{e^{-\lambda r}}{r}. \tag{35}$$

In the representation for which s_1 and s_2 are simultaneously diagonal the total spin is not defined. It is not certain whether the asymptotic values to which the singlet and triplet states tend for $r \to 0$ are the same or not. At any rate, the triplet-singlet difference becomes relatively smaller at small distances and it is possible that it will disappear in an exact solution for $r = 0$.

Although through the higher states all potentials are deepened, no new singularity is introduced in this way. The lower limit (35) decreases as $1/r$ like the ordinary meson potential.

As it is seen from figure 1, the P-P potential is very small at large distances but approaches the P-N potential rapidly for $\lambda r < \frac{1}{2}$. At small distances the *potentials are charge independent*. In our approximation the ratio $f_{\text{P-P}}/f_{\text{P-N}}$ is very nearly equal to unity for $r = 0$, and it becomes exactly equal to unity in a higher approximation. The ratio $f_{\text{P-P}}/f_{\text{P-N}}$ is also plotted in figure 1. Since in our theory the potentials are deeper and more narrow than in the old theory, the scattering of protons or neutrons by protons will, to a larger extent, be due to contributions from small distances. It is quite possible that the scattering experiments can be explained by potentials of the type of figure 1. Whether this is the case or not can only be decided after a more detailed calculation of the scattering cross-sections. Some American authors claim exact identity of the P-P and P-N potentials.* Exact charge independence can clearly not be obtained in our theory and if it should prove to be true we had still to fall back on the introduction of neutral mesons. It is, however, very interesting to note that Hoisington, Share and Breit (1939) from calculations of the scattering cross-sections on grounds of the meson potential $e^{-\lambda r}/r$ came to the conclusion that the range of the singlet

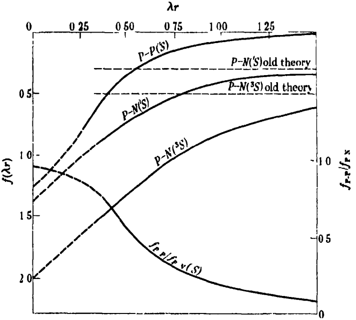


FIGURE 1. P-P and P-N potentials for the ¹S- and ³S-states. The function $f(\lambda r)$ is the potential divided by $e^{-\lambda r}/\lambda r$ (equation (33)). The dotted parts are in a region where the higher approximations are effective. In this region the true curves are lower. The scale on the left-hand side is in units of Λ. One unit corresponds to approximately 20 MeV. The ratio $f_{\text{P P}}/f_{\text{P N}}$ is also plotted (scale on the right-hand side). The constants chosen are: $g^2\lambda = f^2\lambda = 0{\cdot}31$. $f(\lambda r)$, according to the old theory, is constant. The values of f_{old} (dotted) are for the same values of the constants.

---

* We are very much indebted to Professor Breit for a comment on this question.

potentials is only half of that to be expected from the value of the meson mass. This is now exactly what we find in our theory, assuming the right meson mass ($\mu = 180$). This agreement may be considered as a certain support of our hypothesis.

For a more detailed comparison of our results with experiments it must be remembered that the curves of figure 1 are valid for a certain choice of the constants. It is very probable that a slight variation of these constants may lead to a still closer similarity of the P-P and P-N curves. In view of the qualitative and preliminary character of the theory, we have not made any further attempts in this direction.

In connexion with the higher spin states there is another point of a more principal nature which must be borne in mind. In the meson theory—also in the new form proposed by us—a proton has a charge g and a magnetic moment f/λ. A magnetic dipole moment is in fact the highest moment that can be attributed to a particle with spin $\frac{1}{2}$. Particles with spin higher than $\frac{1}{2}$ can, in principle, have *higher moments*. A particle with spin j could have, in addition to an electric monopole and a magnetic dipole, an electric quadripole, a magnetic 8-pole, an electric 2^4-pole, etc., the highest being a 2^{2j}-pole. In our theory particles with spin $\frac{3}{2}$ occur. These particles could have an electric quadripole and a magnetic 8-pole moment and could produce a meson field of such a higher type. If those higher moments existed they would give further contributions to the nuclear potentials, even in the approximation considered above. Whether they exist or not can only be decided on grounds of a more deeply founded theory than that attempted in this paper, but for this reason the curves of figure 1 must be considered with some reserve.*

7. OTHER EFFECTS

In this section we discuss briefly the bearing of our assumptions on other effects which have been derived from the meson theory.

A. *The $1/r^3$ term in the deuteron potential.* In the preceding section we have derived the nuclear potentials for pure S-states. This procedure is not quite exact since the exact deuteron potential (28) leads for the ground state to a coupling of the ^{3}S- with a ^{3}D$_1$-state, the coupling-energy being proportional to $1/r^3$. A $1/r^3$ term inserted in the Schrödinger equation gives rise to a diverging eigenvalue. This difficulty is not remedied by our new assumptions, since the new potential (equation (28)) also contains, of course, the $1/r^3$ term which simply arises from the magnetic dipole-dipole interaction, Møller and Rosenfeld (1939) have tried to avoid this difficulty by assuming that the meson field is a combination of a vector and a pseudoscalar field. Kemmer (1939) has shown recently that there is an intimate mathematical connexion between those two fields and it might well be that this possibility has seriously to be taken into account. Here, we wish to draw attention to another aspect of this problem. At the end of § 6 we mentioned that a particle with higher spin could, in principle, have higher electromagnetic momenta and could produce a meson field of a higher type. Thus the higher spin states might give rise to a further contribution to the interaction of two heavy particles in addition to (28) corresponding to the interaction of two electric quadripoles, or of a magnetic dipole with a magnetic 8-pole and so forth. (28) would only be the first term of an expansion valid for large distances. Although the higher types of interaction would increase with even higher powers of $1/r$, it may well be that the series converges. The interaction of two finite charge distributions may be finite at all distances and yet, if expanded in terms of multipoles, may increase with higher and higher powers of $1/r$. Whether this possibility is to be taken seriously will be a question of the future development of the meson theory.

---

* A meson could thus have an electric quadripole moment. In the ordinary form of the theory it has in fact such a quadripole as a relativistic effect. This can best be seen from Kemmer's formulation of the theory (Kemmer 1939, equations (35), (36)).

B. *Deviation from the Coulomb law.* According to the meson theory in its old form, the electrical field in the neighbourhood of a proton deviates largely from the Coulomb law (Fröhlich, Heitler and Kahn 1939). Owing to the virtual emission of positive mesons, the charge distribution was found to be of the following form: A cloud of positive charge extending over a region with linear dimensions of the order of $1/\lambda$ and increasing rapidly with decreasing distance is compensated by an infinite negative point charge at the position of the heavy particle. Consequently, the Coulomb attraction decreases at small distances and turns into a repulsion for $\lambda r < \frac{1}{2}$. The energy levels of the hydrogen atom are shifted towards higher energies. Whilst earlier experiments seemed to show an effect of this sort, Drinkwater, Richardson and Williams (1940), on grounds of new experiments, came to the conclusion that no real evidence has yet been obtained for it. Probably the effect is smaller than the theory would predict. In our new theory the situation is quite different. A proton is also capable of emitting virtually negative mesons and the result will be that the positive charge density in the immediate neighbourhood of the heavy particle is much smaller than in the former form of the theory. In all probability the repulsive effect will disappear and the law of force will only be a diminished attraction. The shift of the hydrogen energy levels towards higher energies will be very much smaller and will probably lie outside the reach of the present spectroscopical accuracy.

C. *Photoelectric effect of the deuteron.* The exchange forces between proton and neutron contribute considerably to the cross-section for the photoelectric disintegration of the deuteron, especially at high energies. The effect is the larger the more rapidly the nuclear potential decreases with distance. The effect is therefore particularly marked for the meson potential (28), the largest contribution being due to the $1/r^3$ term (Fröhlich *et al.* 1940). This does not mean, however, that it is only the extremely small distances which matter; for an energy of 10 MeV the main contribution comes from distances of the order of magnitude of $\frac{1}{2}\lambda$. In our new theory the deuteron potential decreases even more rapidly with distance. It is therefore to be expected that the large cross-section for the photo-disintegration at high energies derived from the meson theory will remain unaltered in our theory. The effect depends, however, on the presence of the $1/r^3$ part of the nuclear potential. In Møller and Rosenfeld's (1939) modification of the meson theory no such effect would occur. Measurements of the cross-section and angular distribution for the photodisintegration of the deuteron at high energies would provide a very valuable check on the presence of a $1/r^3$ potential and on the two alternative formulations of the meson theory.

D. *Interaction with the electromagnetic field.* In this paper we have only been concerned with the interaction of mesons with nuclear particles. There is, however, little doubt also that the interaction of mesons with the electromagnetic field will have to be changed in some fundamental points. For many effects depending upon this interaction the present theory leads to results which are, at least at high energies, incompatible with experiments. We mention three examples: (i) The cross-section for Bremsstrahlung emitted by a meson during the passage through the Coulomb field of a nucleus is found to increase at high energies like the square of the energy, or, if the modification of the Coulomb field within the nuclear radius is taken into account, like the energy of the meson (Oppenheimer, Snyder and Serber 1940; A. H. Wilson and Booth 1940). The corresponding cross-section for an electron is independent of the energy. The result is incompatible with the high penetrating power mesons with energy $> 2 \times 10^{10}$ eV. The effect is due to the largeness of the interaction between meson and light quantum associated with a change of the direction of polarization. It is interesting that similar large results have been derived for the collision cross-section of a meson with an electron, but these results do not seem to be incompatible with the experiments (Massey and Corben 1939; Oppenheimer *et al.* 1940).

(ii) In a collision with a free nuclear particle a meson can be transformed into a light quantum, for instance,

$$Y^+ + N \rightleftarrows P + h\nu.$$

The cross-section also increases like the square of the energy (Heitler 1938; Kobayasi and Okayama 1939). This is certainly incompatible with experiments for energies higher than a few times 10^8 eV. Although in this process the interaction with the nuclear particle is also involved, the result is not affected by our modification of the meson theory. A closer examination of the calculation shows that the largeness of the effect is solely due to the largeness of the meson-light interaction associated with a change of the direction of polarization of the meson. The same was the case for (i).

(iii) As mentioned in § 6, a meson has a relativistic electric quadrupole moment. A negative meson, say, moving in the Coulomb field of a nucleus has therefore an attractive potential proportional to $1/r^3$ which gives rise to an infinite energy level for the ground state.*

In view of these results it must be concluded that the description of the electromagnetic properties of mesons by the present theory cannot be correct and that some fundamental changes are required in order to bring it into harmony with experiments.

8. THE OBSERVABILITY OF THE NEW PARTICLES

Most suitable for observation are the particles in higher charge states $2e$ or $-e$. The excitation energy being of the order of magnitude of 20 MeV, it is clear that these particles could not have been observed in laboratory experiments. The only possibility of finding them is in nuclear collisions of cosmic ray particles. Roughly speaking, those collisions give rise to two different sorts of processes: (i) The primary energy is taken up by the nucleus and transformed into nuclear heat energy. Even if the primary particle has an energy of 10^3 MeV, the nucleus reaches a temperature of not more than 10–20 MeV. During the subsequent evaporation it is unlikely that particles in one of the excited states will be emitted, although, of course, this may happen in a small fraction of cases.

(ii) The primary energy is transferred to a single nuclear particle which leaves the nucleus at once, taking up an energy comparable with the primary energy. In this case single protons with energies of the order of magnitude of 10^8–10^9 eV are produced and it is to be expected that an appreciable fraction will be in excited states. Again, the chance for observing them at sea level (or even on high mountains) is diminished by two factors. Undoubtedly most of the heavy particles are produced in the high atmosphere. The particles with charge $2e$ lose four times more energy than the particles with charge e and have therefore a much smaller chance than ordinary protons to penetrate through the atmosphere. (This would not apply to particles with charge $-e$.) Secondly, the excited particles are certainly liable to β-decay during which they fall down to the normal proton or neutron state. It is difficult to say what the decay period would be for an energy of 20 MeV. It is certainly a small fraction of a second and it is quite conceivable that the life time is smaller than 10^{-4} sec. From an extrapolation of Sargent's law for light nuclei one arrives at a figure of the order of 10^{-3} sec. This figure may be too large, as has been pointed out by Nordheim and Yost (1937) because the matrix element for the β-transition is much smaller for a complex nucleus than for a single particle. If the life time is 10^{-4} sec. or less, a large fraction of the excited particles decay before they reach sea level. The best chance for finding them is therefore at great heights.

---

* According to a private communication by Kemmer.

So far few heavily ionizing particles have been observed in cloud chamber experiments, and in very few cases (most of them are slow mesons) their identity has been established. It is improbable that the new particles could have been discovered in any of the experiments carried out up to the present.

There is, however, one track published by Anderson and Neddermeyer (1936, figure 13) which might be interpreted as a particle with charge $2e$. In order to identify a track completely three independent measurements (for mass, charge and energy) are required, for instance, range, ionization and curvature in a magnetic field. If for charge and mass only values are assumed which correspond to known particles, two measurements are sufficient. For convenience we give the formulae for range R, ionization I and $H\rho$ for a particle with mass M (large compared with electronic mass) and charge z expressed as a function of the energy E, velocity v and $H\rho$, assuming that the energy is small compared with Mc^2.

$$R = \frac{E^2}{2AMz^2}\frac{1}{\log(aE/M)} = \frac{Mv^4}{8Az^2}\frac{1}{\log(av^2/2)} = \frac{z^2(H\rho)^4}{8AM^3}\frac{1}{\log(az^2(H\rho)^2/2M^2)},$$

$$I = 2B\frac{z^2}{v^2}\log\frac{av^2}{2} = Bz^2\frac{M}{E}\log\frac{aE}{M} = B\frac{M^2}{(H\rho)^2}\log\frac{az^2(H\rho)^2}{2M^2},$$

$$eH\rho = \frac{Mv}{z} = \frac{\sqrt{(2ME)}}{z}. \tag{36}$$

A, B, a are constants which depend only upon the material traversed. For normal air a is about equal to 12 if all masses and energies are expressed in units of the proton mass and MeV respectively. The formulae for R are correct if $2\log(aE/M) \gg 1$, which is approximately the case for the energies in which we are interested.

The track in question emerges from a lead plate and is produced by a big shower. According to Anderson and Neddermeyer it has an $H\rho$ of 'at most' $1\cdot4 \times 10^5$ gauss cm. and a visible range certainly larger than 5 cm. Apparently, it passes out of the illumination, but judging from the photograph published by Anderson *et al.* the curvature seems to increase at the end of the visible part of the track and if this is the case the actual range cannot be much larger than 5 cm. If the track is interpreted as a proton, its energy would be at most 1 MeV and its range would not be more than 2 cm. Next it would be natural to assume that it is a slow meson. Then the range would be, according to (36), 200 m., the energy 10 MeV, and the specific ionization only about thee times greater than the minimum ionization. The appearance of the track does not make this interpretation very likely (compare Corson and Brode 1938), and it is impossible if the increase of curvature near the end of the range is real.

The information supplied by the photograph is in full agreement with the assumption that it is a particle with protonic mass and charge $2e$. According to (36) the range would then just be 6 cm., or 5·4 cm. if the increase of M by the excitation energy is taken into account. The ionization would practically be the same as that of an ordinary proton of the same $H\rho$. Any other explanation would be far less probable. If we would for instance try to interpret the track as due to a particle with smaller mass, the mass would be 1300 m., which is a very unplausible figure.

We give this interpretation with greatest reserve. Before it can be accepted other reasons for the anomalous behaviour of the track, perhaps in the nature of scattering by the gas, will have to be discussed. It must also be remembered that 'unusual' tracks have been occasionally reported by several authors (one of them, for instance, by Williams and Pickup (1938) with an apparent mass of 540 m.) which cannot readily be interpreted as due to particles in higher charge states. Certainly, far more experimental material is required before the question of the existence of the new particles can be settled.

We wish to express our thanks to Dr Fröhlich for many interesting and helpful discussions. Our thanks are also due to Drs Bhabha and Kemmer who have helped to clarify the difficulties which were the starting point for this paper.

SUMMARY

The present meson theory exhibits a number of serious difficulties if applied to the interaction of fast mesons with nuclear particles. In order to avoid these difficulties the following hypothesis is made: Charge and spin of a proton and neutron shall be capable of assuming higher quantum states. Thus particles shall exist with approximately protonic mass and with charges $+2e$, $-e$, etc., and also with spins $\frac{3}{2}$, $\frac{5}{2}$, ... It is assumed that the lowest of these higher states have excitation energies which turn out to be of the order of magnitude of 20 MeV. The higher spin states are introduced by means of a simple model, namely that of a rotator with half-integer angular momentum. The theory is essentially non-relativistic with respect to the heavy particle.

The theory leads to the following results:

(i) The cross-section for scattering of mesons by a heavy particle is smaller by an order of magnitude than according to the former theory (and in qualitative agreement with experiments) and no longer increases with energy. Thus no multiple showers are to be expected.

(ii) The anomalous magnetic moments of the proton and neutron diverge only logarithmically and, if the theory is limited to energies smaller than the rest energy of the proton, are in reasonable agreement with the experimental values.

(iii) The influence of the higher states on the nuclear potentials results in a deepening of the potentials at small distances. Even without introducing neutral mesons there is strong attractive proton-proton potential, and it is discussed whether the proton-proton scattering data can be explained without this hypothetical particle.

(iv) The deviation from the Coulomb law for a proton derived from the former meson theory is much smaller in the new theory, whereas for the photodisintegration of the deuteron the characteristic meson effects are expected to be even more marked.

The possibilities for observing the new particles are discussed. It is pointed out that it would have been unlikely for the particles to have been discovered in any of the experiments which have been carried out up to the present. Nevertheless, one track published by Anderson and Neddermeyer which cannot be due to any of the known particles is tentatively interpreted as due to a particle with charge $2e$ and protonic mass.

REFERENCES

Alvarez and Bloch 1940 *Phys Rev.* **57**, 111.
Anderson and Neddermeyer 1936 *Phys. Rev.* **50**, 270.
Bethe 1940 *Phys. Rev.* **57**, 260, 390
Bhabha 1938 *Proc. Roy. Soc.* A, **166**, 501.
— 1939 *Proc. Roy. Soc.* A, **172**, 384.
Born and Jordan 1930 *Elementare Quantenmechanik.* Berlin.
Corson and Brode 1938 *Phys. Rev.* **53**, 777.
Dirac 1936 *Proc. Roy Soc.* A, **155**, 447
Drinkwater, Richardson and Williams 1940 *Proc. Roy. Soc.* A, **174**, 164.
Fierz and Pauli 1939 *Proc. Roy. Soc.* A, **173**, 211
Frohlich, Heitler and Kahn 1939 *Proc. Roy. Soc.* A, **171**, 269.
— – 1940 *Proc Roy Soc.* A, **174**, 85
Frohlich, Heitler and Kemmer 1938 *Proc. Roy. Soc.* A, **166**, 154.
Heisenberg 1939 *Z Phys* **113**, 61.
Heitler 1938 *Proc Roy. Soc.* A, **166**, 529
— 1939 *Report of the Solvay Conference*
— 1940 *Nature, Lond.* **145**, 29
Hoisington, Share and Breit 1939 *Phys Rev.* **56**, 884
Kobayasi and Okayama 1939 *Proc. Phys.-Math. Soc. Japan*, **21**, 1
Lovell 1939 *Proc. Roy. Soc.* A, **172**, 583.
Kemmer 1938 *Proc Camb. Phil. Soc* **34**, 354.
— 1939 *Proc. Roy. Soc.* A, **173**, 91
Massey and Corben 1939 *Proc. Camb. Phil. Soc.* **35**, 463.
Møller and Rosenfeld 1939 *Nature, Lond.*, **143**, 241 and **144**, 476
Nordheim and Yost 1937 *Phys Rev* **51**, 942
Nordheim and Teller 1938 *Phys. Rev.* **53**, 937.
Oppenheimer, Snyder and Serber 1940 *Phys Rev.* **57**, 75
Williams and Pickup 1938 *Nature, Lond*, **141**, 684
Wilson, A. H. and Booth 1940 *Nature, Lond.*, **145**, 103.
Wilson, J. G. 1940 *Proc. Roy Soc.* A, **174**, 73
Yukawa and Sakata 1937 *Proc Phys.-Math Soc Japan*, **19**, 1084

X-ray structure and elastic strains in copper

BY S. L. SMITH, D.SC., A.C.G.I., *Engineering Department*,
AND W. A. WOOD, D.SC., *Physics Department*,
National Physical Laboratory, Teddington, Middlesex

(*Communicated by G. W C. Kaye, F.R.S.—Received* 18 *June* 1940)

(Plates 10–13)

An X-ray examination of copper has been undertaken whilst the metal has been actually under tensile stress, and the X-ray structure investigated at a systematic series of points on the load-extension curve and during unloading and reloading from selected points on the curve. It is shown that the permanent strain is associated with the breakdown of the grains into the crystallite formation and that this change is essentially irreversible The elastic strain of the metal is accompanied by reversible changes in dimensions of the atomic lattice which take place without leaving in the lattice any permanent distortion, as shown by the observation that the X-ray diffraction rings, including rings of the diffuse type, contract and expand under the action of the applied stress without any change in the degree of diffusion. The lattice changes are distinguished in this way from certain lattice strains or lattice distortions permanently imposed on the lattice as a result of deformation by cold-work. Quantitative measurements are made on the elastic lattice strains exhibited by the (400) and (331) planes in a direction perpendicular to the axis of the applied stress, and these are compared with the equivalent external elastic constants. These measurements show that marked difference in rate of strain may take place in neighbouring grains subjected to the same external stress, and on this difference is based an explanation of the extensive breakdown of the grains into components of widely varying orientations which characterizes the structure of a polycrystalline metal after deformation beyond the yield point.

1. INTRODUCTION

When a metal is stressed beyond the yield point, the resulting deformation may be regarded as consisting of two types, an elastic strain and a superimposed permanent strain. The changes in crystalline structure associated with the permanent strain can be studied without the necessity of examining the material whilst actually under load, since this deformation remains after removal of the stress. This procedure has in fact been followed in recent researches, in which the changes in structure produced by permanent deformation at normal temperatures have been investigated by X-ray diffraction methods; a metal has been subjected to systematic increases in permanent strain by application of appropriate stresses, and the load then removed at each stage whilst X-ray tests were made. By this means, information has been obtained, from the X-ray point of view, on the mechanism by which the metallic grain accommodates itself to this type of strain; in particular, it has been shown that within the grain exists a large-scale unit of structure, termed crystallite, defined by a lower limiting size which has a characteristic value for a given metal; and, also, results have been obtained on the relation between the dispersal of the grains into this crystallite formation and such mechanical properties as yield-point, plasticity and fracture (Wood 1939; Wood 1940; Wood and Thorpe 1940; Gough and Wood 1936).

The above procedure gave no information on the changes associated with the elastic strain, since this effect vanishes as the applied stress is removed. The present research, however, represents a fresh step in which this drawback is overcome by the use of an experimental arrangement in which X-ray examination can be made whilst a specimen is under load and the behaviour under X-rays compared with simultaneous mechanical measurements of the applied stress and the elastic strain. At the same time, it was realized that, whilst the relation between X-ray structure and elastic strain was the main consideration of the work, the picture would not be complete unless the accompanying effects of permanent deformation were also taken into account, since the two types of strain are not entirely unconnected. Thus, as the permanent strain imposed on a metal is increased, then, in general, the possible range of elastic strain is also increased. Some relation is therefore probable between the structural changes underlying elastic properties and the process of breakdown and crystallite formation associated with permanent strain. Both aspects of deformation were consequently studied at the same time by the X-ray method.

2. Scope of tests

The straining apparatus was designed to apply direct tension to flat tensile specimens of the shape and dimensions shown in figure 1.

The X-ray tests are described most conveniently by reference to the imaginary load-extension diagram depicted in figure 2.

The tests were of two types

(i) X-ray examination was made at regular intervals of stress as a specimen was loaded progressively from the initial state to fracture. Thus, with reference to figure 2, the specimen was examined at the points

A_1, B_2, C_3, D_4, ..., lying on the load-extension curve. The specimen was therefore under load throughout the run of a test, and the two processes of elastic strain and permanent strain were taking place simultaneously.

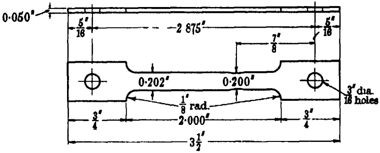


FIGURE 1. Shape and dimensions of X-ray tensile specimen.

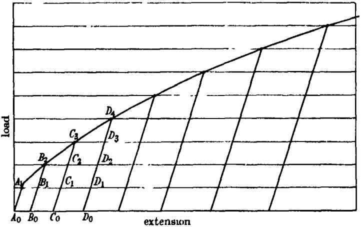


FIGURE 2. Diagrammatic representation of scope of X-ray tests.

(ii) X-ray examinations were also made as a specimen was taken through a series of cycles of loading and unloading, the range of each cycle being increased in regular steps until fracture. The points of examination are perhaps made clearer by reference to figure 2. Thus a specimen was loaded to the stress corresponding to A_1 on the load-extension diagram and there photographed The specimen was then unloaded, the load-extension curve following the approximately straight line A_1A_0 as elastic contraction occurred, and a further photograph obtained. The specimen was next loaded to B_2 along the path which, to a first approximation, is represented by $A_0A_1B_2$, and photographs were obtained at B_2 and, during subsequent unloading, at B_1 and B_0. The process was then repeated and photographs secured at C_3, C_2, C_1 and C_0, and so on. The specimen is thus taken successively through the cycles $A_1A_0A_1$, $B_2B_0B_2$, $C_3C_0C_3$, ..., in each of which, except possibly for second-order effects, elastic changes only are taking place, although the degree of superimposed permanent strain, represented by the distances of A_0, B_0, C_0, ... from the origin, has a different value for each of the cycles. The point of these experiments was therefore that a comparison of the X-ray photographs obtained in a given cycle would indicate the changes in structure associated with elastic strain alone, whilst comparison of the behaviour in different cycles would reveal any additional effects due to differences in degree of superimposed permanent strain Confirmatory tests were made on a number of specimens, and, in addition, a particular cycle such as $C_3C_0C_3$ was examined in detail by taking photographs both when ascending and descending, in order to find whether differences in structure occurred at corresponding loads in the two halves of the cycle.

During each cycle extensometer measurements were taken of the external elastic change in length, measured over a gauge length of 2 in. on the parallel test portion of the specimen. These measurements were thus directly comparable with the X-ray observations.

3. EXPERIMENTAL

The material used for the present experiments was copper of specially high purity, which was selected as representative of a ductile metal of the face-centred cubic type. Each specimen received a preliminary annealing *in vacuo* to remove the effects of machining and to bring the grains into a work-free condition. This was shown by the nature of the initial X-ray photographs, which then exhibited distinct sharp reflexion spots arising from individual grains. This initial condition has been made a special feature of the X-ray work; it ensures that subsequent observations are not confused by preliminary residual distortions of structure, and also it allows of an X-ray technique whereby the behaviour of the same recognizable group of grains can be followed through important stages of a test.

The load-extension diagram obtained separately on standard tensile test pieces of circular section is shown in figure 3 (*a*). The yield occurred at 2–2·5 tons/sq. in. and fracture at about 15 tons/sq. in. nominal stress. The value of Young's modulus was $6{\cdot}6 \times 10^3$ tons/sq. in., and Poisson's ratio, specially measured, was found to be 0·3.

The tensile-straining apparatus consisted essentially of a miniature testing machine by means of which known loads could be applied to the specimen. Briefly, the load was applied by weights on a scale pan attached to a piston floating in a cylinder of oil, and the pressure transmitted by way of a flexible pipe to a second part of the apparatus where the pressure forced upwards a piston to which the free end of the specimen was linked. This part of the apparatus which carried the specimen was compactly designed and was rigidly mounted on an X-ray back-reflexion spectrometer developed for these researches. The surface of the specimen under examination could be brought into line with the axis of the spectrometer, and, owing to the flexible nature of the transmission pipe, the specimen could be photographed when oscillating about the spectrometer axis as well as when stationary. A further feature of the apparatus was that the mount carrying the specimen could be adjusted so that, as the specimen extended on loading, the incident X-ray beam could be directed always on to the same point of the specimen, a feature to which particular importance is attached since it ensures, for instance, that any changes observed will not be caused by local differences in the behaviour of grains situated at different places in the specimen. Finally, a sensitive mirror extensometer was attached to the specimen and arranged so that the changes in its elastic strain during the cycles of loading and unloading could be read off from the deflexion of a spot of light on a scale.

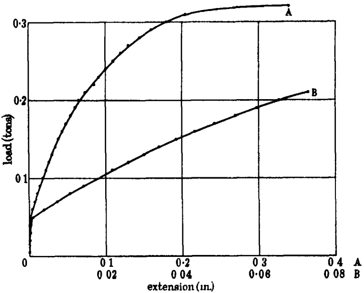


FIGURE 3. Load-extension diagram of copper. (*a*) Complete curve. (*b*) Initial portion of curve to larger scale. Cross-sectional area 0·0195 sq. in., gauge length 1·0 in.

The X-ray back-reflexion photographic method was used because the recorded diffraction rings, being formed by deviation through very large angles, are highly sensitive both to lattice changes and also to the dispersal of the grains into variously oriented crystallites, whilst at the same time least sensitive to possible experimental variations in the distance between specimen and film.

Cobalt radiation was employed for the incident beam, and the back reflexions were recorded as complete rings symmetrical about the direct beam. These rings were the $\alpha_1\alpha_2$ doublet from the (400) planes and the β ring from the (331) planes. The time of exposure was of the order of 3–5 min

A point to note is that the selective action of any X-ray method based on the photography of an aggregate assumes special importance in an investigation of the present type, since each diffraction ring is formed by a different set of grains. In the usual application to analysis of atomic structure, this point has no particular significance, as it is known that the atomic arrangement is the same in each grain, in the present work, however, the behaviour of a grain under stress is likely to vary according to its orientation, because of the anisotropic nature of the atomic structure. It is therefore necessary to note the crystallographic orientation of the grains which take part in the formation of the particular diffraction rings under observation. The incident beam is perpendicular to the surface of the flat tensile specimen in its stationary position, and the effective (400) planes are those for which the Bragg reflexion angle is $\theta = 81{\cdot}7^\circ$, that is, the grains concerned are those in which a cube face makes an angle of $\frac{1}{2}\pi - \theta = 8{\cdot}3^\circ$ with the surface of the specimen, or, to a first approximation, those in which the cube face lies parallel to the surface. Similarly the (331) ring refers to those grains in which the (331) planes make an angle of 12·4° with the surface of the specimen. The X-ray method is thus of interest because even in a random aggregate it permits examination of grains of selected orientation; by tilting the specimen at various angles to the beam, the behaviour of grains lying at any orientation to the applied stress could be explored. It is hoped to extend the work along these lines when experimental arrangements of the required precision are developed; the present results however refer to the grains oriented in the manner described above.

4. RESULTS

In the first tests under continued loading, X-ray examination was made at intervals of 0·5, 1, 1·5, 2, ... tons/sq. in. (on the lines described in § 2), the specimen being held at each stress for about 5 min. whilst the X-ray photograph was obtained. As the fracture stage was approached, this time interval was extended until the creep was negligible during the period of exposure. The observations showed two types of structural change; first, the large-scale modifications in structure as the grains were broken down into the complex of fragments and crystallites, and, second, a superimposed and progressive contraction of the (400) and (331) lattice spacings, which indicated a quite distinct fine-scale effect associated with the dimensions of the atomic lattice. It was, of course, to be expected that if lattice changes occurred they would be in the direction of contraction, since, as pointed out in the preceding section, the effective grains are those in which the (400) and (331) planes are inclined at small angles to the surface of the specimen, so that the measured spacing of each family of planes is roughly perpendicular to the direction of the applied tension. It was rather surprising, however, to find that the observed contraction should be the same for all such grains, independently of the way the cube face lay in the surface of the specimen, this was shown by the fact that the diffraction rings remained completely circular throughout, the contraction due to change in the lattice spacing being the same along each radius.

In the second class of tests, as described in § 2, the cycles employed were respectively 2 to 0, 4 to 0, 6 to 0, ... tons/sq. in. The comparison of the X-ray photographs obtained in each cycle showed only the lattice effect; the lattice contracted under load, and at once recovered upon removal of the load by an amount proportional to the change in elastic strain. Since in these experiments the diffraction ring is not affected by further plastic deformation in each cycle, the changes in diameter due to the lattice variations could be obtained with greater accuracy than in the preceding tests. It was found that there was marked difference in the behaviour of the (400) and (331) planes.

It follows that (*a*) the process of plastic deformation is concerned primarily only with the irreversible transition of the grains into the crystallite formation, whilst (*b*) elastic strain is associated with the reversible expansion and contraction of the atomic lattice in the fragments and crystallites constituting this formation. It is therefore convenient first to consider the two processes (*a*) and (*b*) separately, and then finally (*c*), the relation between them.

(*a*) *The permanent changes*

The changes associated with plastic deformation may be dismissed fairly briefly since they are essentially similar under load to those described for other metals in previous papers (*loc. cit.*), where the earlier procedure was adopted of making X-ray tests only after applied stresses had been removed. For the purpose of discussion of the further results it is, however, desired to draw attention to three distinct phases exhibited by the crystallite formation. These are the following:

(i) *The initial phase*, extending up to the yield-point, where the original grains preserve their structure as a whole, the X-rays indicating no appreciable differences in orientation of the component parts of each The X-ray photograph of this phase is characteristic; the grains give separated reflexion spots, which exhibit no tendency to spread along the circumference of the diffraction rings, as they would if the crystallographic orientation within each grain were not uniform A typical photograph is reproduced in figure 4 (plate 10); the slight radial spread of the spots is due to the oscillation of the specimen through ± 5° about the spectrometer axis, when the grains reflect over a small angular range during the oscillation at the large diffraction angles involved.

(ii) *The plastic phase*, which sets in as soon as the elastic range is passed and the transition to which marks the yield-point. Here the grains form a heterogeneous mixture of fragments and crystallites with orientations considerably different from the parent grains. The X-ray photograph is again characteristic, there is a spread of reflexions along the circumference of the diffraction ring producing a ring which becomes continuous but is of irregular intensity distribution. The type is shown by figures 5–10 (plates 10–13).

(iii) *The fracture phase*, into which the preceding structure passes as the plastic deformation increases and the specimen begins to neck. A random distribution of crystallites then preponderates in the structure and the lower limiting size of these units is approached, a value which for copper is about $0{\cdot}7 \times 10^{-5}$ cm. (*loc. cit.*). The effect on the X-ray photograph is to produce a continuous and uniform distribution of intensity round the diffraction rings and, since the lower limiting crystallite size is too small for perfect resolution (Scherrer effect), the rings become diffuse, the (400) doublet broadening into a single halo. The resulting type of photograph is shown in figure 11 (plate 13).

In each of these phases, therefore, there is no difference as the specimen is unloaded, despite the lattice change which causes the diffraction ring to expand as a whole. This point is of special interest in the fracture phase when the ring is very diffuse; the fact that the ring contracts or expands bodily without effect on the degree of diffusion, although the internal lattice strain is varying considerably, confirms the view that this present line-broadening is due to the Scherrer effect and not to irregularities in spacings arising from some other form of internal strain.

(b) *The elastic changes*

The elastic changes are first illustrated by typical X-ray photographs to show the order of the effect to be measured, and the quantitative results then summarized.

(i) For convenience of illustration the photograph of the diffraction ring obtained under a given load has been cut along a diameter and matched against the corresponding half of the photograph obtained after removal of the load. One side of the (400) rings is made to coincide and the displacement at the other side gives the change in diameter. Since the external elastic strain of the specimen as a whole is small compared with the area illuminated by the X-ray beam, the same grains or particles are reflecting throughout a cycle; the rings, therefore, preserve the same appearance through each cycle. This permits accurate determination of the changes in diameter in the initial and plastic phases, where the intensity is irregularly distributed round the rings, and also in the fracture phase where the rings are diffuse since measurements can always be made on the same selected diameter in each photograph. The photographs are reproduced in figures 6–11. They suffice to show that the changes are of an easily measurable order under the experimental conditions employed.

TABLE 1. EFFECT OF STRESS ON DIAMETER OF (400) DIFFRACTION RING

Cycle (ton/sq. in.)	Diameter of (400) ring in mm. at following stresses (ton/sq. in.)								Change in diameter per ton/sq. in. mm.
	0	2	4	6	8	10	12	14	
2 to 0	39·8	39·5							0·15
4 to 0	39·8	39·5	39·2						0·15
6 to 0	39·8	39·5	39·3	39·0					0·13
8 to 0	39·8	.	39·2	39·0	38·7				0·14
10 to 0	39·7	39·5	39·2	38·9	38·6	38·3			0·14
12 to 0	39·7	.	39·1	.	38·5	.	37·9		0·14
14 to 0	39·7	.	.	.	.	.	.	37·6	0·15
								Mean	0·14

(ii) The quantitative results obtained on a typical specimen taken through stress cycles of increasing range are shown in table 1. Measurements on other specimens were the same within experimental error. This table gives the observed diameter of the (400) ring for various loads within each cycle, and indicates that to a first approximation the change in

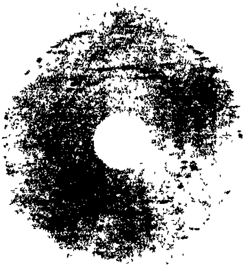


FIGURE 4. Initial phase.

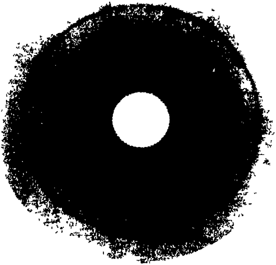


FIGURE 5. Underload 2 tons/sq in, showing change to plastic phase of X-ray structure on passing through yield point.

(Facing p. 406)

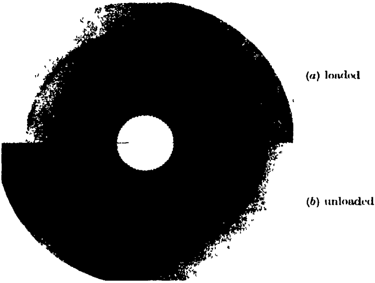


FIGURE 6 6 tons/sq. in.

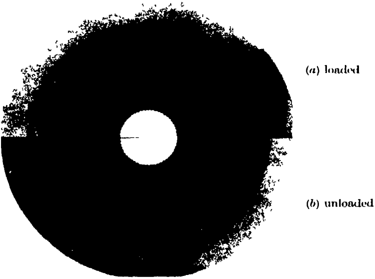


FIGURE 7. 8 tons/sq in

The top half (*a*) of each photograph refers to the loaded, and the bottom half (*b*) to the unloaded condition. Each pair is so mounted as to bring the left sides of the inner (400) doublet ring into coincidence; the displacement of the other side shows the change in diameter produced by unloading.

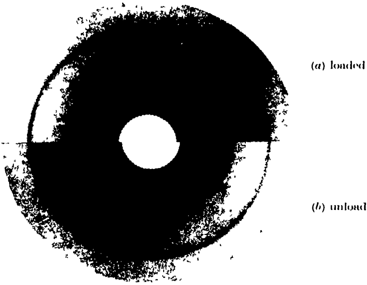


FIGURE 8 10 tons/sq. in.

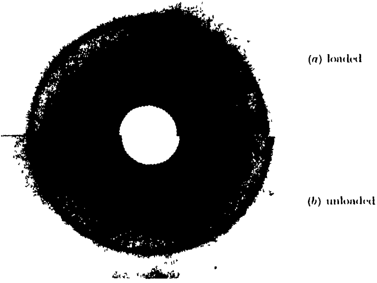


FIGURE 9 12 tons/sq. in.

Figures 8 and 9 illustrate increased change in diameter of (400) ring on unloading from higher stresses

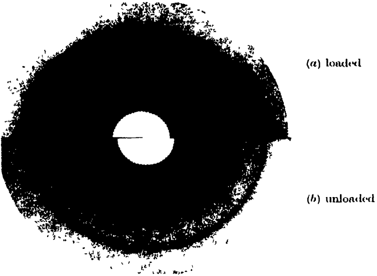


FIGURE 10. 14 tons/sq in

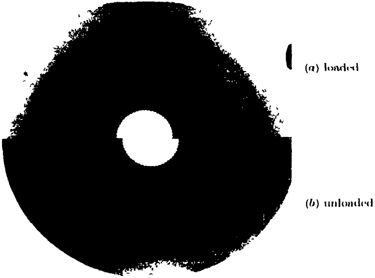


FIGURE 11 Fracture phase.

Figures 10 and 11 illustrate further contraction of (400) diameter with stress Figure 11, secured just before fracture (c 15 tons/sq. in) shows also diffuse rings characteristic of fraction phase under static stress diameter bears the same proportional relation to the applied stress in each cycle, and is therefore independent of the effects of the superimposed permanent strain which is different for the various cycles.

The lattice spacing of the (400) planes was calculated in the usual way for a series of assumed diameters of the diffraction ring and was found to be directly proportional to the diameter over the range of variation. The change in spacing which corresponded to a change in ring diameter of 1 mm. was then obtained from the same calculations. The measurements in table 1 may be summarized by the contraction in diameter per unit increase in stress, and this quantity, since the relation between ring diameter and spacing was known, was then expressed directly as the percentage contraction in (400) lattice spacing per unit stress (ton/sq. in.) tension along the axis of the specimen. As indicated above, this contraction is in the direction making an angle of 8·3° with the normal to the applied stress. Since this angle is small, it is reasonable to obtain the value of the contraction in the direction actually perpendicular to the applied tension by dividing the observed value by cos 8·3°. The results are summarized below together with the corresponding values obtained in the same way from measurements on the (331) planes. Also the elastic changes in external dimensions as given by the extensometer measurements are included for purpose of comparison.

(*a*) *Lattice spacings of* (400) *and* (331) *planes*

		(400)	(331)
(1)	Spacing change corresponding to 1 mm. change in diameter of diffraction ring (calibration constant) (A)	0·00045	0·00056
(2)	Diameter change produced per unit stress (ton/sq. in.) applied along specimen axis (mm.)	0·14	0·04
(3)	Spacing change per unit stress (from 1 and 2) in directions 81·7 and 77·6° respectively with applied stress (%)	−0·0069	−0·0027
(4)	Spacing change per unit stress in direction perpendicular to applied stress (%)	−0·0070	−0·0028

(*b*) *External extension and contraction of specimen*

(5)	Extension in direction of applied stress per unit stress (ton/sq. in.) (%)	0·015
(6)	Contraction perpendicular to applied stress per unit stress (%)	−0·0045

The contraction in (6) was obtained from (5) by application of the measured value of Poisson's ratio of 0·3. The lattice contractions given in (4) are thus fundamental physical constants expressing in terms of internal atomic displacements the externally measured contractions in the same direction. It will be noted that whilst the observed external contraction of 0·0045 % must result from the value of the lattice contractions averaged out over the various grains in the cross-section of the specimen, the individual percentage contractions for differently oriented grains may vary greatly, thus the (400) contracts by 0·0070 %, whilst the (331) spacings contract only by 0·0028 %. The percentage contraction of these spacings, both of which refer to the same direction, namely, that perpendicular to the applied stress, therefore differs by as much as 2·5 to 1 under the same change of stress. It may be indicated that, apart from the calculations of the final values, this difference was evident from direct inspection of the X-ray photographs. The sensitivity of the (400) and (331) rings to a given percentage change in spacing is not very different, namely, as 56 to 45, but the photographs showed at once that the (400) ring was roughly three times as sensitive as the (331) to the same applied stress. This point is mentioned because some importance is attached to the extent of the difference in connexion with its possible bearing on the changes produced by permanent deformation. It will be noted also that the total contraction over a particular grain is proportional to the appropriate percentage lattice contraction multiplied by the width of the grain so that, since the grains may vary appreciably in size, the actual difference in contraction of contiguous grains of different orientation may in some instances be considerably greater than the differences indicated by the figures for the percentage lattice contraction. The observations thus raise interesting issues concerning the behaviour of the structure at the boundaries of such grains, a point discussed below.

(c) *The combination of the two types of observations*, the effects of permanent and of elastic strain, yields a more complete picture of the response of the crystalline structure of a metal to static stress. The main feature of the permanent strain is the striking dispersal of the grains into the crystallite formation, and it is reasonable to conclude that there is a connexion between this instability of the structure of the grains and the effects of the elastic strains. Thus, up to the yield-point, in the state described as the initial phase, the original grains are observed to remain individually stable in that they exhibit no appreciable amount of the breakdown; but at the yield-point an instability sets in, and differently oriented components are extensively formed. The instability at the yield must presumably arise either because the degree of the elastic strain in the individual grains has reached a critical value, or as a consequence of the interference between the different amounts of elastic strain exhibited by neighbouring grains of different orientations in the manner indicated above. It is suggested that the effective factor is the second, and that the form taken by the observed instability, namely, the dispersal of the grains into differently oriented components, is the mechanism by which potential local discontinuities in strain between such neighbouring grains are smoothed out by transition through the appropriate newly formed orientations. This view is supported by the further observation that the new structure is then stable over an increased range of lattice strain; the atomic lattice of the various fragments or crystallites can therefore take up a greater variation than the change produced by the yield stress, when other conditions are favourable, without disturbing the overall stability of these components. Similarly, as deformation is continued, a further increase of stress causes more fragmentation and variations in orientation, whilst at each state a correspondingly greater range of elastic lattice strain is then permissible without producing instability in the new units thus formed. When, however, the fracture phase is approached, a fresh factor is introduced, since the X-ray observations show that the lower limiting size of the crystallites is reached so that the process of accommodation to increasing stress by continued breakdown must become more difficult and a more critical condition must ensue.

It is of interest to note the value of the lattice strain at the yield and at fracture. For the (400) planes, or the edge of the unit cubic cell, the lattice change as measured becomes 0·014 % for a specimen which yields at 2 tons/sq. in. and as much as 0·11 % for the breaking stress of 15–16 tons/sq. in The corresponding values for the (331) planes are respectively 0·006 and 0·04 %.

5. Discussion

It is of interest to distinguish between the elastic lattice strain investigated in this paper and the lattice distortion studied in recent work on the more severe deformation of a metal by cold-work (Wood 1939). The present results show that over the whole range represented by the load-extension diagram it is possible in copper to impose on the atoms a reversible displacement from their normal positions which leaves no measurable distortion in the lattice; this is shown by the fact that in the stress cycles employed a reversible change of diameter took place without any accompanying change in diffusion of the X-ray reflexions in any of the three phases, initial, plastic or fracture. On the other hand it was shown, in the work referred to, that a change in diameter of the diffraction rings could also be produced which, however, persisted after the process of cold-working; in this instance, moreover, the change was accompanied by a marked abnormal diffusion of the rings, and when the ring diameter was reduced to its normal value by further treatment of a specimen then the additional diffusion was also removed. This lattice strain, on account of the extra diffusion, was shown to be a permanent strain associated with distortion of the lattice. The comparison between the two types of results shows that it is possible in this way to distinguish between a purely elastic lattice strain and this permanent lattice distortion. The distinction becomes necessary when attempts are made to estimate the equivalent stresses in a metal from X-ray measurements of the amount of lattice strain, a point which arises in connexion with the problem of internal strains. It is known that as a result of certain treatments a metal may be left in an unstable condition, recovery from which may lead in the course of time to changes of shape or in extreme cases to fracture. The metal is said to be in a state of internal strain, and attempts have been made to estimate the equivalent stresses from X-ray measurements of the permanent lattice changes combined with use of the elastic moduli. It is, however, clearly essential that the nature of the lattice strain should be known, since it is only in the case of the elastic strain, together with data on the lines obtained in the present experiments, that the method would be justifiable. The internal strain of the lattice distortion type, which is more likely to be involved in internal strains, would entail a more detailed knowledge than is at present available both of the distortion itself and of the interatomic forces in the lattice.

One of the puzzling features of the X-ray structure of a ductile polycrystalline metal after deformation has been the extent of the breakdown of the grains into the independently reflecting units of widely varying orientations. The effect is much more marked than in the early stages of deformation of a large single crystal. It has already been suggested that the process is a grain-boundary effect in that the breakdown arises from the necessity of preserving continuity of structure across the boundaries of grains oriented in such a way that the rate of strain is different on each side of the boundary. The fragmented structure is then stable, apart from possible second order changes in the X-ray structure, for ranges of stress up to a limit depending on the extent of the fragmentation That such second-order changes are, however, often present is probable, for it is unlikely, if an applied stress cycle were sufficiently large or if it approached the limit for a given crystallite formation, that the adjustment of orientations would be perfect enough on the first cycle to compensate completely for the variation in elastic strain; the adjustments might also be expected to vary slightly during subsequent repetition of the cycles as well as with the speed of application of the cycle. The point is mentioned because it would account, at least in part, for the departure from a strictly linear stress/strain relationship which is in practice often observed during application of a stress cycle, and for subsequent variations in width of the hysteresis loop thus formed, which may also be observed on repeated application of a cyclic stress. The sensitivity of the X-ray method to detect minor changes in the stability of the structure depends largely on the nature of the X-ray photograph The conditions are most favourable when the diffraction ring consists initially of sharp spots, as used in the recent work on fatigue stressing when it was in fact found that slight but continued changes could always be detected when the applied stress cycle exceeded the fatigue limit for the material. The present results now suggest an explanation of the changes found in that work.

In conclusion, the authors wish to express their thanks to Dr G. W. C. Kaye, F R.S., for his continued interest in the work.

REFERENCES

Gough, H. J. and Wood, W. A. 1936 *Proc Roy. Soc.* A, **154**, 510.
Wood, W. A. 1939 *Proc. Roy. Soc.* A, **172**, 231.
— 1940 *Proc Phys Soc* **52**, 110.
Wood, W. A. and Thorpe, P. L. 1940 *Proc Roy. Soc.* A, **174**, 310.

The influence of the solvent on the formation of micelles in colloidal electrolytes

I. Electrical conductivities of sodium dodecyl sulphate in ethyl alcohol-water mixtures

By A. F H. Ward

College of Technology, University of Manchester

(*Communicated by E. K. Rideal, F.R S.—Received* 10 *July* 1940)

In very dilute solutions in water sodium dodecyl sulphate behaves as a completely dissociated electrolyte. Above a concentration $c = 0\ 00722$ N the equivalent conductance (Λ) falls sharply and micelles are formed. To throw light on the mechanism of micelle formation, the variation of Λ with c has been measured in a series of mixtures of water and ethyl alcohol. The curves for Λ against $\sqrt{c}$ show a gradual transition from the type associated with micelle formation to a uniform curve characteristic of strong electrolytes.

Addition of alcohol decreases the tendency to form micelles. When there is 40 % or more by weight of alcohol, micelles are no longer formed. The critical concentration first falls on addition of alcohol and then rises again. The effects of alcohol addition can be interpreted satisfactorily by considering the radius of the micelle to be equal to the length of the paraffin chain and thus to be independent of the concentration and nature of the solvent.

The energy changes involved in micelle formation are calculated. Aggregation can occur if the interfacial energy available from the destruction of the paraffin-solvent interfaces of the ions is greater than the work to be done against electrical repulsion The interfacial energy is no longer the greater beyond 40 % of alcohol. It is shown that alcohol molecules are not in solution in the paraffin interior of the micelles in the mixed solvents, but are strongly adsorbed on the micelle surfaces.

The influence of regions of lowered dielectric constant around the micelles is considered. This is more important at high concentrations. The final fall in Λ with highly supersaturated solutions is attributed to the increased viscosity of the solution.

An aqueous solution of a salt, one ion of which contains a long paraffin chain, shows a normal electrical conductivity at low concentrations. That is, the equivalent conductance, when plotted against the square root of the concentration, gives a straight line, in agreement with the Onsager equation. Above a critical concentration, there is a sharp fall in the conductance. This is interpreted as being caused by micelle formation (Moilliet, Collie, Robinson and Hartley 1935). McBain (1913) originally postulated the formation of micelles to explain the *rise* in conductance of soap solutions at higher concentrations, since in such an aggregate of ions the increased charge carried would more than counterbalance the increased viscous resistance to motion. If micelle formation is to explain a *fall* in conductance, the smaller ions of opposite charge (gegenions) must be assumed to be carried with the micelle in close association (Hartley 1935), reducing its effective charge and decreasing the number of free gegenions. In agreement with this, several independent methods show without doubt that micelles are present above the critical concentration. Measurements of the conductances of several colloidal electrolytes with high-frequency current (Schmid and Larsen 1938) show that highly charged ions are present above the critical concentrations. Similarly, the size of the particles calculated from diffusion coefficients (Hartley and Runnicles 1938), and the solubility of organic solvents in the interior of the micelles (Hartley 1938) show, in the case of cetyl pyridinium salts, that micelles are present above the critical concentration. The value of the critical concentration for micelle formation depends on several factors. It is decreased by an increase in the length of the paraffin chain, by a fall in the temperature or by added electrolytes. It may also be altered in the presence of organic substances dissolved in the water.

The object of the present investigation was to examine in detail the mechanism by which the solvent influences micelle formation. The gradual alteration in behaviour was observed on passing from aqueous solutions in which micelles are formed to alcoholic solutions where there are no micelles The formation of micelles was examined by measuring the equivalent conductance of the solutions Sodium dodecyl sulphate was chosen as the solute. The absence of hydrolysis with this salt makes possible accurate measurement even in great dilution. Conductance measurements in aqueous solutions have been made by Lottermoser and Puschel (1933) and by Howell and Robinson (1936). The present author has already measured conductances in mixtures of water and ethyl alcohol with very low concentrations of the salt (Ward 1939). These measurements are now extended to the highest attainable concentrations, in some cases supersaturated solutions. The maximum concentration reached with the aqueous solutions was fifteen times that of a saturated solution.

EXPERIMENTAL

Apparatus

The electrical circuit was the same as that previously used by the author. Alternating current with a frequency of 2000 cycles was given by a valve oscillator. A Tinsley non-inductive resistance box and Tinsley ratio arms formed the bridge, and similar ratio arms were used for the Wagner earth. A variable condenser in parallel with the resistance box allowed complete silence to be obtained at the null point. This was determined with a valve amplifier and telephones.

Three quartz cells with lightly platinized electrodes were used. The cell constants were measured with several solutions of Kahlbaum's potassium chloride. The specific conductances of these were of the same order as those of the solutions for which the cells were to be used. The values of the constants were 0·7403, 0·02912 and 0·01355. A sufficient volume of solution was always used in the cell so that the cell constant no longer varied with the volume.

Materials

The sodium dodecyl sulphate was prepared by the action of chlorosulphonic acid on dodecyl alcohol and the neutralization of the resulting acid with caustic soda. The salt was purified by crystallizing from water (twice) and from alcohol (five times), and by extracting with petroleum ether for 5 hr. The temperature was always kept below 60° C to avoid decomposition. An analysis gave Na $7\cdot97_3$ %, hydrolysable SO_3 $27\cdot6_4$ % (theoretical values 7·979 and 27·77 %).

Conductivity water of specific conductance about $0\cdot5\times10^{-6}$ reciprocal ohm was obtained by fractional condensation in a silver condenser. It was stored in silica vessels. Absolute alcohol was purified by distillation from lime. The middle fraction was collected. A little of this was refluxed with calcium shavings for 1 day, and to the resulting gel of calcium ethoxide the rest of the alcohol was added. This was refluxed for another day and then the alcohol was distilled into a silica bottle. This specimen was extremely pure and had a specific conductance of only $0\cdot05\times10^{-6}$ reciprocal ohm.

The alcohol-water mixtures were made by diluting redistilled rectified spirits with conductivity water. The specific gravities were measured and the compositions found from tables. Table 1 gives the values of the percentage of ethyl alcohol by weight and the corresponding specific conductances.

These mixtures will be referred to by the nearest whole number in the percentage of alcohol.

Procedure

To measure the specific conductances of the more concentrated solutions in the small cell, the salt and the solvent were weighed directly in the cell.

The solution was diluted either by simply adding more solvent or by first withdrawing solution and then adding solvent by the Ostwald-Arrhenius method. All withdrawals or additions were determined by weighing. For the more dilute solutions in the larger cells the pure solvent was first weighed in the cell and its conductance measured. Then known weights of a solution of suitable concentration were added from a weight pipette of borosilicate glass. This method, due originally to Whetham (1900), gives the concentration and the solvent correction more definitely than does the dilution method. For solutions in absolute alcohol, the solubility was too small to allow the concentration to be increased by adding a more concentrated solution from the weight pipette. Consequently, suitable additions of the solid were added to the solution in the cell.

TABLE 1

Alcohol % by weight	Specific conductance	Alcohol % by weight	Specific conductance
9·91	0 33 × 10⁻⁶ mho	59 37	0·11 × 10⁻⁶ mh
20·02	0 21	69·23	0·12
29 78	0·16	80 15	0·12
39 60	0·12	93·04	0·10
49 94	0 12		

All measurements were made at 20 ± 0·01°.

The concentrations were known originally in weight normalities. Since comparisons must be made between solutions in solvents of different densities it is necessary to convert the concentrations to volume normalities. Below concentrations of about 0·01 N the density of the solution can be taken as being the same as that of the solvent, with a maximum error of 2 parts per 1000. But for the more concentrated solutions examined here, the densities differ appreciably from those of the pure solvents and must be known. Density measurements were therefore made with solutions in each of the solvents, over the whole concentration ranges investigated. These will be the subject of a later paper, but the values obtained are used here in the calculation of equivalent conductances.

RESULTS AND DISCUSSION

In table 2 the values of the equivalent conductance (Λ) at 20° are given for various values of $\sqrt{c}$ (c = concentration in g.-equivalents per litre). The values for very low concentrations (from 0·0001 to about 0·01 N) have already been published (Ward 1939) and are not included. In order to

TABLE 2

Water		9·91 % alcohol		20·02 % alcohol		29·78 % alcohol	
$\sqrt{c}$	Λ	$\sqrt{c}$	Λ	$\sqrt{c}$	Λ	$\sqrt{c}$	Λ
0·07403	61·56	0·07347	45·12	0·09292	34·44	0·08414	27·35
0·08143	61·20	0·07679	44·58	0·1050	33·70	0·1103	27·11
0·08835	60·68	0·08172	43·74	0·1234	32·59	0·1203	26·97
0·09363	59·03	0·09226	41·10	0·1454	31·14	0·1343	26·65
0·09665	56·40	0·1320	34·41	0·1553	30·42	0·1450	26·40
0·1228	43·38	0·1655	30·67	0·1676	29·47	0·1892	25·33
0·1608	34·02	0·2006	28·30	0·2188	26·85	0·2451	24·13
0·2041	29·68	0·2420	26·65	0·2436	26·05	0·3059	23·14
0·2332	28·10	0·3087	25·36	0·3195	24·46	0·3991	21·92
0·2784	26·89	0·3613	25·05	0·4166	23·83	0·4912	21·21
0·3180	26·38	0·4402	25·44	0·4775	23·68	0·5332	20·98
0·3687	26·26	0·4742	25·74	0·5929	23·77	0·6339	20·63
0·4224	26·46	0·5193	26·11	0·7532	23·43	0·7761	19·54
0·4777	26·95	0·5832	26·73	0·8958	22·50	0·8884	18·76
0·5468	27·75	0·6313	27·20	1·031	21·17	1·033	17·30
0·6146	28·60	0·7119	27·70				
0·7063	29·61	0·7836	27·76				
0·7771	30·02	0·8474	27·58				
0·8612	30·04	0·9656	26·38				
0·9330	29·51	1·108	23·47				
0·9913	28·52						
1·052	27·10						
1·120	25·23						

39·60 % alcohol		49·94 % alcohol		59·37 % alcohol		69·23 % alcohol	
$\sqrt{c}$	Λ	$\sqrt{c}$	Λ	$\sqrt{c}$	Λ	$\sqrt{c}$	Λ
0·1050	23·87	0·09658	22·59	0·1248	21·17	0·1286	20·74
0·1279	23·52	0·1225	22·00	0·1579	20·45	0·1622	19·69
0·1563	23·12	0·1533	21·45	0·1954	19·65	0·2003	18·62
0·1873	22·59	0·1896	20·72	0·2308	18·90	0·2363	17·63
0·2201	22·19	0·2296	20·03	0·2974	17·66	0·3054	16·17
0·2701	21·50	0·2728	19·24	0·3850	16·27	0·3788	14·82
0·3555	20·46	0·3251	18·49	0·5009	14·62	0·5053	12·76
0·4590	19·29	0·3806	17·80	0·6491	12·75	0·6628	10·49
0·5675	18·36	0·4124	17·42	0·8407	10·53		
0·7302	16·81	0·4792	16·60				
0·8900	15·33	0·5937	15·30				
1·119	12·84	0·6955	14·28				
		0·8254	12·93				
		0·9454	11·69				
		1·044	10·68				

80·15 % alcohol		93·04 % alcohol		Absolute alcohol	
$\sqrt{c}$	Λ	$\sqrt{c}$	Λ	$\sqrt{c}$	Λ
0·1146	21·15	0·1266	19·44	0·1156	17·38
0·1432	19·80	0·1648	17·18	0·1320	15·92
0·1882	18·19	0·2131	14·93	0·1543	14·52
0·2389	16·63	0·2591	13·37	0·1770	13·29
0·3042	14·96	0·2973	12·30		
0·3852	13·17				
0·4682	11·52				

save space, many values, agreeing with those given in the table, are omitted.

Values of Λ are plotted against $\sqrt{c}$ in figure 1. Many experimental points for low concentrations are omitted to avoid confusion. The curves are shown as broken lines where the solutions are supersaturated.

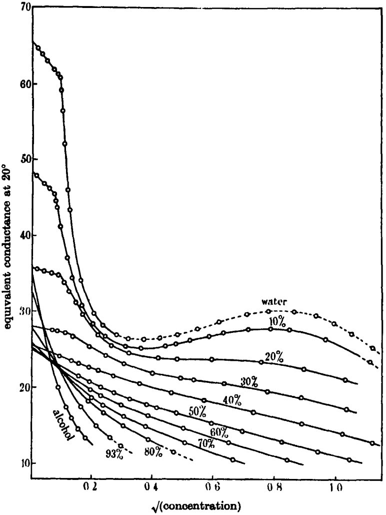


FIGURE 1. - - - - supersaturated solution.

The curve for aqueous solutions can be divided into four regions. (1) an initial straight line, where the equivalent conductance falls in accordance with the Onsager equation (up to $\sqrt{c} = 0{\cdot}085$); (2) a very rapid fall, reaching a minimum at $\sqrt{c} = 0{\cdot}36$, (3) a more gradual rise, to a maximum at $\sqrt{c} = 0{\cdot}84$; followed by (4) another fall, increasing in slope with increasing concentration.

When alcohol-water mixtures are used as solvents, lower values are obtained for the equivalent conductance. The values at infinite dilution (Λ_0) first fall with increasing alcohol content (minimum at 50 % of alcohol) and then rise again. This is because of the maximum in the viscosity-concentration curve for alcohol-water mixtures. As the alcohol content rises, there is also a decrease followed by an increase in the initial slope of the linear portion of the curve. This has been interpreted by the author in terms of the Onsager theory, applied to unassociated ions.

What is more important in considering micelle formation is the alteration in the character of the curves. As alcohol is added, the rapid fall in Λ at the critical concentration becomes progressively less marked. With 30 % of alcohol the break in the curve is barely detectable and at 40 % of alcohol it no longer occurs. Simultaneously the maximum in the curve at higher salt concentrations becomes flatter when alcohol is added. With 20 % of alcohol the minimum and maximum have very nearly disappeared, and with 30 % of alcohol there is merely a slight decrease followed by an increase in the slope of the curve. The curves from 40 % of alcohol onwards show no unusual features. After the initial linear portions, the slopes steadily decrease with increasing salt concentration in the manner usual with concentrated salt solutions.

Sharp fall in equivalent conductance

According to Stokes's law, one would expect micelle formation to cause a rise in the equivalent conductance Hartley (1935) explained how micelle formation could cause a fall in the equivalent conductance. The retardation of ions caused by the Coulomb forces ('atmosphere' effect) must increase with the charge. The mobilities will be further decreased by the inclusion of gegenions with the micelle ('inclusion' effect), thus decreasing both the resultant charge on the micelle and also the number of free gegenions. These two secondary effects tending to diminish Λ will conceal the expected increase and actually cause a fall at the critical concentration.

The manner in which these secondary effects are influenced by adding alcohol to the solvent can be considered. Addition of alcohol gives a gradual lowering of the dielectric constant and a relatively larger increase in the viscosity, up to about 40 % of alcohol (see figure 2, curves 4 and 5). Onsager's theory cannot be applied quantitatively to very highly charged ions, but it is easy to see from Onsager's equation in which direction these alterations in dielectric constant and viscosity affect the atmosphere effect. The decrease in dielectric constant causes an increase in the atmosphere effect. The increase in viscosity causes a much bigger decrease in the effect, both by its direct action and also indirectly by decreasing Λ_0. On the whole, therefore, the diminution in Λ caused by the atmosphere effect should be smaller as alcohol is added to the solvent.

The inclusion of gegenions on to the surface of a micelle will be affected chiefly by the dielectric constant. A fall in the dielectric constant, by the addition of alcohol, should cause the gegenions to be attracted more strongly to the micelle, thus giving a bigger decrease in Λ. (An increased tendency towards ion association is generally found when the dielectric constant is decreased.)

Thus, the resultant alteration in Λ when micelles are formed at the critical concentration in alcohol mixtures will be made up from a primary increase (Stokes effect) diminished by the atmosphere and inclusion effects. The atmosphere effect will be smaller than with water and the inclusion effect larger. Now it is found experimentally that from 40% of alcohol onwards there is no critical concentration, and the $\Lambda-\sqrt{c}$ curves are smooth. This must mean that with each mixture containing more than 40% of alcohol *either* there are no micelles formed *or* the increases and decreases in Λ caused by the above-mentioned effects exactly cancel each other. The latter explanation is so improbable that it can be said that the effect of increasing the alcohol content of the solvent mixture is that the fraction of ions which aggregate to micelles at the critical concentration is gradually decreased and falls to zero at about 40% of alcohol. As the number of micelles is thus diminished, the magnitude of the various effects produced by micelles must simultaneously become smaller, giving a progressively smaller drop in Λ at higher alcohol concentrations.

Critical concentration

At the critical concentration there is equilibrium between single ions and micelles. Grindley and Bury (1929) have shown by application of the law of mass action that an appreciable increase in the number of micelles may be expected over a small increase of total concentration. If the addition of alcohol to the solutions displaces the equilibrium in the direction of single ions, one might expect that the critical concentration would be higher. Table 3 gives the values obtained for the critical concentration

The expected increase in the critical concentration is found in the range from 10 to 30% of alcohol, but there is a fall in going from water to 10%

of alcohol. However, a closer analysis of the mechanism of micelle formation explains this apparent anomaly.

TABLE 3

	$\surd c$	c
Water	0·085	0 00722
10 % alcohol	0·072	0·00518
20 % alcohol	0·085	0·00722
30 % alcohol	0·102	0·0104

Micelle formation

The aggregation of ions occurs because they expose less hydrocarbon surface to the water when they are grouped as micelles. A spherical micelle would not be expected to have a radius greater than the length of the paraffin chain, as otherwise some polar heads would be drawn into the non-polar interior. The arguments of Hartley and Runnicles (1938) lead them to conclude that 'there will be a fairly well-defined optimum size for the micelles *almost independent of concentration*, and that this size will correspond to a sphere of radius of the order of the length of a fully extended chain'. Their experimental results from diffusion measurements agree with this conclusion.

The length of the paraffin chain in sodium dodecyl sulphate is 15·2 A, and this can therefore be taken as the radius of the micelle. It will have a surface of 2900 A^2 and a volume of 14,700 A^3. The number of aggregating ions will be limited by the number of paraffin chains which can fit into this volume. The volume of one molecule of dodecane at 20° C is 366 A^3, giving 40 ions per micelle. It is possible to calculate the decrease in interfacial energy when the aggregation occurs, using the method developed by Langmuir (1926). The interface between water and the paraffin chains of the ions will disappear, and when these chains are in contact in the micelle there will be no interfacial energy between them. The interfacial energy associated with the polar groups will not change, since these are surrounded by water both before and after micelle formation.

The total surface energy between water and a hydrocarbon of high molecular weight is 59 ergs per cm.^{2}. The area of the paraffin chain of a single ion (assumed to be rolled up in the water approximately like a sphere) is 248 A^2. The original interfacial energy per ion is therefore $59 \times 248 \times 10^{-16} = 146 \times 10^{-14}$ erg. The part of the surface of the micelle covered with polar groups can be taken as 825 A^2 (allowing 20·5 A^2 for each polar group), so that the micelle still exposes a hydrocarbon surface of $2900 - 825 = 2075$ A^2 to the water. The total surface energy of this, calculated *per ion*, is 30×10^{-14} erg. The fall in total surface energy on micelle formation is therefore $146 \times 10^{-14} - 30 \times 10^{-14} = 116 \times 10^{-14}$ erg per ion This can be compared with the value of $kT = 4 \times 10^{-14}$ erg.

If the interfacial energy were the only work involved in micelle formation, the Boltzmann equation would give the ratio of the number of ions united in micelles to the number of free ions $= Be^{116/4} = 3{\cdot}9 \times 10^{12} B$ (where B is a constant involving the *a priori* probabilities of the two states). If this were so one might expect that even at great dilutions free ions would not exist. However, during micelle formation work must be done in overcoming the electrical repulsion between the ions. This cannot be calculated very definitely. The inclusion of gegenions, forming an electrical double layer round the micelle, would decrease the electrical work. The order of magnitude can be found by calculating the energy of charging of a sphere the size of the micelle with 40 electronic charges, neglecting the effect of the positive ions and assuming that the negative charges were originally separated infinitely. If the radius of the sphere of paraffin chains in the micelle is 15·2 A, it is reasonable to allow an extra 3 A for the polar groups, giving 18·2 A as the radius of the charged sphere. If the micelle is surrounded by water of dielectric constant 80 the calculated energy of charging is $31{\cdot}6 \times 10^{-14}$ erg *per ion*. This is of the same order as the interfacial energy. It is impossible to say exactly how much the positive ions will cause this value for the electrical energy to be diminished, without knowing the number and disposition of the included gegenions To take an extreme case, if a number of gegenions, equal to the number of negative ions, formed a layer at a distance 3 A farther from the centre than that of the negative ions, the electrical energy would be decreased to $4{\cdot}6 \times 10^{-14}$ erg per ion.

There is another effect which probably causes a big *increase* in the electrical energy. The value of the dielectric constant used in the above calculations is that of water in bulk. The value is smaller in the immediate neighbourhood of charged ions, owing to the orientation of solvent molecules (Williams 1931) This has been indicated from conductivities of sodium dodecyl sulphate at greater dilutions (Ward 1939). Sack (1927) has calculated the lowering of dielectric constant near a univalent ion in water, the value being very small within 3 or 4 A and gradually rising to 80 at about 15 A from the ion. This 'pocket' of lower dielectric constant occupies a greater volume the higher the valency of the ion, the radius being multiplied by $\sqrt{z}$ for a z-valent ion. The size of the ion, as distinct from its charge, has an inappreciable effect. On the assumption that this is still true for a micelle, with 40 charges, the region of lower dielectric constant would extend to about 95 A from the micelle. At the critical concentration in water the distance between ions of the same sign is 61 A. If all the negative ions were to be aggregated to micelles at this concentration, their separation would be 209 A. Whether or not this estimate of the size of the region of lowered dielectric constant can be relied on accurately, it is clear that at, and above, the critical concentration it must be an appreciable fraction of the volume of the solvent near the micelles. Thus the true value of the electrical energy in micelle formation must be considerably greater than the approximate value estimated above, and may be nearer to the value of the interfacial energy.

It is probable that the values for interfacial and electrical energy are not very different. Below the critical concentration the energy that would be obtained from the destruction of the paraffin-water interfaces is insufficient to overcome the electrical repulsions. Above the critical concentration, the ions are not so distant from each other initially and less work has to be done against the electrical forces in forming micelles. The interfacial energy is now sufficient to do this. Supporting evidence for this view is given below in considering the effects of alcohol addition. It is known that the addition of even small concentrations of ions, particularly multivalent ions, to the solution has a considerable effect in promoting micelle formation (Powney and Addison 1937).

Micelle formation in presence of alcohol

Since the size of the micelle is controlled fundamentally by the length of the paraffin chain it should be the same in water-alcohol mixtures as in water. However, as is shown below, the charge may be different. The main difference in the micelle caused by the presence of alcohol in the solvent is that some alcohol molecules may be included in the micelle. This could occur in two ways: (1) dissolved in the paraffin interior, (2) adsorbed along the paraffin-water interface of the micelle.

No data appeared to be available on the solubility of alcohol in higher paraffins from alcohol-water mixtures. Experiments were therefore made by shaking medicinal paraffin (freed from polar impurities with Fuller's earth) with two alcohol-water mixtures containing 10 and 30% of alcohol respectively. The alcohol which dissolved in the paraffin was extracted with water and estimated colorimetrically with concentrated sulphuric acid and potassium dichromate. Values for the partition coefficient from the two solutions were in agreement.

$$\frac{\text{Concentration of alcohol in paraffin}}{\text{Concentration of alcohol in water}} = 0{\cdot}0044$$

(concentrations being expressed per unit volume) at 20° C. The amount of alcohol dissolved in a single micelle from the 10 % alcohol mixture can be calculated from this to be 0·085 molecule. It is clear, therefore, that the solubility of alcohol in the interior of the micelles can be neglected.

To find how much alcohol is adsorbed on the surface of the micelles, and to calculate the change in interfacial energy on micelle formation, it is desirable to know how the interfacial tension between a hydrocarbon surface and an alcohol-water mixture depends on the percentage of alcohol. Since this information was not already available, the interfacial tension was measured at 20° C by the method of Harkins and Brown (1919) between medicinal paraffin and various alcohol-water mixtures. The results are shown in figure 2, curve 1.

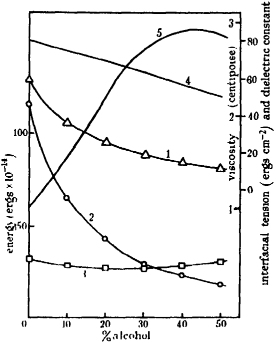


FIGURE 2. Variation with percentage alcohol of interfacial energy (curve 2) and electrical energy (curve 3) (scale on left), interfacial tension (curve 1), dielectric constant (curve 4) and viscosity (curve 5) (scales on right).

The interfacial tension first falls sharply as oriented alcohol molecules are adsorbed in the interface. With high alcohol concentrations, the interfacial tension does not alter much since the interface is nearly packed with alcohol molecules. The fraction of the interface occupied by alcohol can be estimated sufficiently accurately for the present purposes by taking it to be proportional to the fall in interfacial tension. These figures conform with the more rigid calculations of Schofield and Rideal (1925) on the constitution of the interface between air and alcohol-water mixtures.

When 10% of alcohol is added to the water, the interfacial tension is decreased by 41%, i e a fraction of 0·41 of the interface between the solvent and paraffin can be taken to be occupied by alcohol molecules. If there are n ions in the micelle, the paraffin surface is $(2900-20{\cdot}5n)$ A^2. Taking 21·6 A^2 as the cross-sectional area of the polar head of the alcohol, the number of alcohol molecules adsorbed on the surface of the micelle

$$=\frac{(2900-20{\cdot}5n)\times 0{\cdot}41}{21{\cdot}6}.$$

The volume of the C_2H_5 groups of these, buried in the paraffin of the micelle,

$$=\frac{(2900-20{\cdot}5n)\times 0{\cdot}41\times 61{\cdot}3}{21{\cdot}6}\,\text{A}^3 = 1{\cdot}16(2900-20{\cdot}5n)\,\text{A}^3.$$

Since the total volume of the micelle is the same as in water (14,700 A^3), the volume available to be filled by the paraffin chains of the ions $= 14{,}700 - 1{\cdot}16(2900-20{\cdot}5n)$ A^3 This must equal $363\times n$, the volume of n paraffin chains. Solving this equation gives the number of chains $n = 33$, and the number of alcohol molecules in the surface is 42.

In calculating the interfacial energy made available by micelle formation, the total surface energy should be used rather than the free surface energy or interfacial tension, as given in figure 2. However, the observed interfacial tension between medicinal paraffin and water (58·2 erg/cm ^{2}) is not very different from the accepted value of 59 erg/cm.^{2} for the total interfacial energy between a higher paraffin and water, and the values with alcohol-water mixtures may also be assumed to be similar in the two cases. When account is taken of the original interfacial energy of the ions and the interfacial energy still possessed by the micelle, an energy of 65×10^{-14} erg per ion is available on micelle formation. This is a little more than half the corresponding energy with water as solvent.

Similar calculations have been made for other solvent mixtures and the results are shown in table 4

The electrical energy in charging a sphere the size of the micelle is also given in each case for comparison. This is calculated as with water, taking the bulk value of the dielectric constant of the solvent. Figure 2 shows how the interfacial energy and the electrical energy vary with alcohol concentration (curves 2 and 3) There is a large fall in the interfacial energy, but relatively little change in the electrical energy, since the lower dielectric constant compensates partially for the smaller charge on the micelle.

TABLE 4

	Water	Percentage alcohol				
		10	20	30	40	50
No. of paraffin chains in micelle	40	33	30	27	26	25
No. of alcohol molecules in surface of micelle	—	42	64	79	86	93
Interfacial energy per ion (erg $\times 10^{-14}$)	116	65	43	29	23	18
Electrical energy per ion (erg $\times 10^{-14}$)	31·6	27·9	26·9	27·1	28·4	30·6

Thus, micelle formation should be possible in those solvent mixtures where the value of the interfacial energy is greater than that of the electrical energy. Because of the various assumptions made and the many doubtful factors which may affect the electrical energy, the point of intersection of the two curves cannot be taken as giving quantitatively the limiting concentration of alcohol for micelle formation. It is probably a coincidence that it gives a concentration between 30 and 40 % of alcohol, in good agreement with the experimental results. But qualitatively, it is clear that the assumptions made must result in the interfacial energy falling more rapidly than the electrical energy and giving *some* alcohol concentration beyond which micelles cannot be formed.

The decrease in critical concentration on going from water to 10 % alcohol means that the first addition of alcohol causes micelles to be formed more readily This is probably because the alcohol adsorbed on the paraffin surface of the micelle causes a big fall in the interfacial energy and makes the micelle more stable. This lowering in energy is much greater for the first addition of alcohol than for subsequent additions, because alcohol is adsorbed into the interface quite strongly from dilute solutions of alcohol in water. Once this alcohol is on the paraffin surface, subsequent additions make much less difference to the character of the interface than the original alteration from an interface with no alcohol at all.

Rise in equivalent conductance

Less is known about the rise in Λ at higher concentrations than about the fall at the critical concentration. At high concentrations, where it is difficult to interpret the electrical behaviour of even simple salts, it is impossible to say exactly what effects would be caused by highly charged micelles. There is no reason to suppose that gegenions would be desorbed from the micelles in this concentration region, since they are assumed to be adsorbed in more dilute solutions.

An effect which possibly contributes to this rise follows from the conception of pockets of lower dielectric constant around an ion or micelle. Suppose that the dielectric constant in the neighbourhood of a 40-valent micelle in water is very low indeed up to about 25 A ($=\sqrt{40} \times 4$ A) and rises to a value of 80 about 95 A from the micelle. The distances between micelles, calculated on the approximation that all negative ions are aggregated, would be 209 A at $\sqrt{c} = 0{\cdot}085$ (critical concentration), 80 A at $\sqrt{c} = 0{\cdot}36$ (minimum), 46 A at $\sqrt{c} = 0{\cdot}84$ (maximum) and 36 A at $\sqrt{c} = 1{\cdot}2$ (highest concentration reached). As the concentration increases, the pockets of lower dielectric constant from neighbouring micelles will overlap, causing a greater lowering of dielectric constant. This gives a bigger fall in Λ than would be calculated using the bulk value of the dielectric constant, since both atmosphere and inclusion effects are enhanced. However, when the dielectric constant of a part of the solvent near an ion is at the smallest possible value, overlapping with a similar part of the field of another ion can cause no further decrease. Thus, at higher concentrations and small distances between the micelles, the decrease in Λ from atmosphere and inclusion effects is relatively not so great as at low concentrations. Hartley (1939) has suggested the smoothing of the irregularities in the concentration of gegenions to explain the rise in Λ. Now the gegenions would tend to be concentrated in the pockets of lower dielectric constant, and consequently when these overlap at high concentrations, the irregularities of distribution of the gegenions would decrease. This phenomenon would therefore make more pronounced the effect suggested by Hartley.

Final fall in equivalent conductance

With alcohol-rich solvents, the sodium dodecyl sulphate crystallized very readily if the concentration exceeded the saturation value. In mixtures containing much water it was very easy to obtain supersaturated solutions. With water, the solution was actually saturated at a concentration lower than that with a minimum value of Λ. Measurements were made up to a concentration fifteen times that of saturation. In this region Λ was falling quite rapidly with concentration. It is probable that this was caused by the large increase in the viscosity of the solution. The more concentrated solutions were like syrup, a normal solution having a viscosity of about twenty times that of water. The viscosity increase is quite sufficient to explain the final fall in Λ.

REFERENCES

Grindley, J. and Bury, C. R. 1929 *J. Chem. Soc.* pp. 679–684.
Harkins, W. D. and Brown, F. E. 1919 *J. Amer. Chem. Soc.* **41**, 499–524.
Hartley, G. S. 1935 *Trans. Faraday Soc.* **31**, 31–50.
— 1938 *J. Chem. Soc.* pp. 1968–1975.
— 1939 *Kolloidzschr.* **88**, 22–40.
Hartley, G. S. and Runnicles, D. F. 1938 *Proc. Roy. Soc.* A, **168**, 420–439.
Howell, O. R. and Robinson, H. G. B. 1936 *Proc. Roy. Soc.* A, **155**, 386–406.
Langmuir, I. 1926 Alexander's *Colloid Chemistry*, **1**, 525–546.
Lottermoser, A. and Puschel, F. 1933 *Kolloidzschr.* **63**, 175–192.
McBain, J. W. 1913 *Trans. Faraday Soc.* **9**, 99–101.
Moilliet, J. L., Collie, B., Robinson, C. and Hartley, G. S. 1935 *Trans. Faraday Soc.* **31**, 120–129.
Powney, J. and Addison, C. C. 1937 *Trans. Faraday Soc.* **33**, 1253–1260.
Sack, H. 1927 *Phys. Z.* **28**, 199–210.
Schmid, G. and Larsen, E. C. 1938 *Z. Elektrochem.* **44**, 651–658.
Schofield, R. K. and Rideal, E. K. 1925 *Proc. Roy. Soc.* A, **109**, 57–77.
Ward, A. F. H. 1939 *J. Chem. Soc.* pp. 522–530.
Whetham, W. C. D. 1900 *Phil. Trans.* A, **194**, 321–360.
Williams, J. W. 1931 *Chem. Rev.* **8**, 303–319.

The surface as a limiting factor in the slow combustion of hydrocarbons

By R. G. W. Norrish, F.R.S. and J. D. Reagh
Physical Chemistry Laboratory, Cambridge University

(*Received* 17 *May* 1940)

With special reaction vessels, varying widely in diameter while maintaining approximately constant volume, and employing several refinements in technique, a study has been made of the effect of surface on the slow oxidation of several hydrocarbons, both saturated and unsaturated. All reactions were of the degenerate branching type, and were found to be principally homogeneous in character. When the diameter of the reaction vessel was sufficiently reduced, the reaction rate was observed to drop abruptly toward zero, while the corresponding induction period increased toward infinity. Data so obtained have demonstrated that in narrow vessels surface deactivation can predominate over other processes of deactivation, while in wider vessels surface and volume deactivation occur to a comparable extent over a considerable range of pressure When the vessel diameter is decreased to a critical value, the surface deactivation almost alone can suppress the factors leading to chain branching, and from the reactions investigated the existence of such a critical diameter appears to be a general property of hydrocarbon oxidations in conformity with the theory of degenerate branching.

Introduction

One of the most important criteria of branching chain reactions in the gas phase is the existence of critical conditions of temperature, pressure, and diameter of reaction vessel. The oxidation reactions of the hydrocarbons have been shown by Semenoff (1930, 1932, 1935) to take place by a process of branching chains termed degenerate, which he proposed in explanation of the much greater induction periods shown by these reactions than by those of the simpler branching reactions, such as the oxidations of hydrogen and carbon monoxide. According to the theory of branching chains the induction period τ gives a measure of the net branching factor ϕ ($\phi\tau = \text{const.}$, approximately), which controls the development of the reaction, and for a self-accelerating reaction characterized by the equation

$$\omega = Ae^{\phi t} \tag{1}$$

must be greater than zero. The hydrocarbon oxidations, which are controlled then by much smaller values of ϕ than those of the simple branching reactions, require a special mechanism to explain this slower development of the reaction. This mechanism postulates the formation of a comparatively stable intermediate by oxidation of the hydrocarbon through a straight chain reaction, while this intermediate is subsequently oxidized to end products, occasionally in an alternative reaction giving rise to additional chain centres long after the parent chain has been destroyed.

It is a well-established fact that these general principles cover the oxidation of saturated hydrocarbons. The stable intermediate is almost certainly the derived aldehyde, which has been isolated in nearly all cases of the oxidation of hydrocarbons. While the aldehyde may oxidize to end products such as carbon monoxide, carbon dioxide, and water, it may also react with oxygen to yield chain centres, probably oxygen atoms, for instance, as illustrated by the equation

$$\mathrm{RCHO} + \mathrm{O_2} = \mathrm{RCOOH} + \mathrm{O}. \tag{2}$$

Such a mechanism has been worked out in detail for the methane oxidation by Norrish (1935) and by Norrish and Foord (1936), in which the role of intermediate is taken by formaldehyde. The oxidation reactions of the unsaturated hydrocarbons have also been shown to have the characteristics of degenerate chain branching; the work of Bone, Haffner and Rance (1933) on the ethylene oxidation is important in this respect.

It is a matter of experience that self-acceleration towards ignition or fast reaction is dependent on critical conditions of temperature, pressure, and diameter of reaction vessel, all of which affect the magnitude of ϕ. Semenoff has demonstrated that, in chain reactions exhibiting degenerate branching, ϕ can be expressed in the form

$$\phi = k_2(\nu - 1) - k_1, \tag{3}$$

where the constants k_1 and k_2 represent the oxidation of intermediate into end products by alternative reactions, k_2 alone representing the reaction in which additional chain centres are formed, and where ν is the length of the unbranched primary chain By further analysis it is possible to formulate the conditions governing the critical diameter. In a system where straight chains are broken both in the volume (g), and at the surface (S), the total destruction of chains is proportional to β, and the probability of a chain being broken at any particular link is

$$\beta = a(g + S) = a\left(g + \frac{\pi^2 D}{d^2}\right), \tag{4}$$

in which the surface deactivation factor S is expressed in terms of the vessel diameter d, and of the coefficient of diffusion D, a being a constant.

Further, since by definition ν is inversely proportional to β, the following substitution from equation (4) may be made

$$\nu = \frac{1}{\beta} = \frac{1}{a\left(g+\frac{\pi^2 D}{d^2}\right)} = \frac{bd^2}{gd^2+\pi^2 D}. \tag{5}$$

This expression when substituted in equation (3) gives

$$\phi = k_2\left(\frac{bd^2}{gd^2+\pi^2 D}-1\right)-k_1. \tag{6}$$

It is apparent that, for sufficiently large values of diameter, ν becomes constant, and ϕ consequently achieves a constant value dependent solely on the factors which influence the homogeneous production and oxidation of the intermediate substance. As the diameter d decreases, however, the chain length ν becomes increasingly dependent on the diameter, and, *for sufficiently small values of* d, will approximate the form $\nu = cd^2$, where c is a constant. In this case ϕ tends to the form

$$\phi = k_2(cd^2-1)-k_1 = C_1 d^2 - C_2. \tag{7}$$

It is impossible to say precisely in what manner the oxidation of the intermediate is affected by increasing surface (decreasing d), but it may be seen that any change in the factors controlling the destruction of the intermediate which would cause either an increase in C_2 or a decrease in C_1, or both, could bring the net branching factor ϕ to zero or to some negative value while the diameter still remains at finite size. An illustration of the dependence of ϕ on the diameter of the reaction vessel as based on the foregoing theoretical development is shown in figure 1.

The first observation suggesting that the vessel diameter can be a critical limiting factor in chain reactions below the ignition point was made by Spence (1932) Studying the slow oxidation of acetylene by a flow method, he established values for the reaction rate which appeared to show that at a diameter of about 4 mm. the rate dropped away abruptly to very small values; this remaining rate he considered as being caused by a residual surface reaction.

Subsequently Norrish and Foord (1936) investigated the kinetics of the slow oxidation of methane, in the course of which they also observed the existence of a critical diameter below which the reaction was inhibited. This work, using a constant volume method and aided by a number of refinements in technique, showed the critical diameter to be somewhat less than 5 mm., a value of the same order as that obtained by Spence in the acetylene oxidation. Later, Shantarovitch (1937) determined the critical pressure for the oxidation of arsine in vessels of three different diameters, and it is evident that in this reaction too a critical diameter for slow oxidation at a given total pressure must exist. Sadownikow in 1937 concurrently found a similar effect in a study of the critical conditions of the ethane oxidation. Prettre (1938), in work partially contemporaneous with the present investigation, has found the same for *n*-pentane.

It has been the purpose of the present investigation to make a systematic study of the phenomenon of vessel diameter as a critical condition of the oxidation of hydrocarbons, and to show that this limiting factor which was pointed to by the above-mentioned cases may be considered as a general property of degenerate branching reactions.

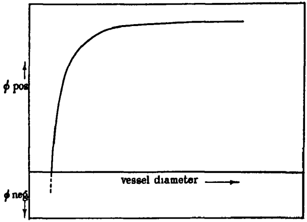


FIGURE 1. Variation of ϕ with diameter of reaction vessel (theoretical development).

Previous work in this laboratory has shown that good reproducibility in experiments in which surface plays a part is obtained if the reaction vessel is kept thermostatic during a series of experiments, at no time being allowed to cool. Many of the irregularities previously recorded have been due to the alternate heating and cooling of the surface, occasionally in the presence of air The accuracy of the constant volume method has been improved by the use of special reaction vessels which were constructed from sections of tubing of any desired diameter joined together in parallel, by this means vessels of approximately constant volume, although varying widely in diameter (3–40 mm.), were obtained Limitations consequent upon the use of the comparatively crude method of packing vessels have thereby been circumvented, as well as the inaccuracies due to loss of volume by reduction of diameter in single-segment vessels. From the results of the present work, which has comprised a study of the slow oxidation of the hydrocarbons methane, ethane, propane, acetylene, ethylene, and propylene, it will be apparent that expectations based on the theory of degenerate branching relative to the surface effect have been fully justified; the diameter has been found in all cases studied to be a critical factor in the sense that observable reactions either cease abruptly or are severely inhibited below a finite diameter.

EXPERIMENTAL METHOD

A. *Apparatus.* The slow oxidation of various hydrocarbons was studied, as described in the following pages, by observation of the pressure change in gas mixtures while at constant volume and specified temperature. An apparatus of a type similar to that used by Norrish and Foord in two previous investigations (1935, 1936) was employed; see figure 2. Essentially it consisted of an electrically heated reaction vessel into which previously prepared gas mixtures could be admitted from a separate mixing vessel, the ensuing pressure change being observed by either of two specially modified Bourdon gauges. Extensions to the manometer, as well as inlets for hydrocarbons and oxygen, are indicated on the sketch, while the mixing vessel A, reaction vessel B, heated entry tube C, and one Bourdon gauge D are indicated by letters alone. The automatically regulated furnace could be kept constant within one degree of any desired temperature in the range 0–700° C. The apparatus could be evacuated to 10^{-5}–10^{-6} mm. of mercury by 'Hyvac' and mercury diffusion pumps.

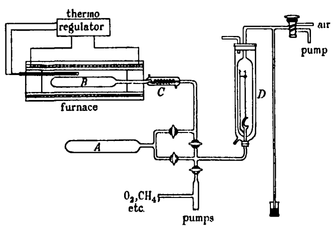


FIGURE 2. Diagram of apparatus.

Of the two Bourdon gauges, one was of a type designed and described by Foord (1934). As well as being strengthened to the point of withstanding one atmosphere on one side of the spoon, it was sufficiently sensitive to be read to 0·05 mm. Hg. Readings were taken from a fixed scale upon which a reflected light beam was cast by a small concave mirror affixed to the pointer of the gauge. The gauge had a linear range corresponding to 120 mm. of mercury and, combining sensitivity with great flexibility, was valuable in determining reaction rates which were very fast, as well as easily accommodating considerable changes in pressure. Throughout these tests this gauge was used for comparatively fast reactions, say 10 mm. Hg/min. and over.

The second gauge, although more sensitive, was more limited in its use. Pressure changes were determined by the movements of its long pointer as observed through a travelling microscope. Although it could be read to 0·02 mm. Hg, its linear range was not great, and manual operation of the microscope necessarily implied less speed of observation. Consequently this gauge was used for the slower reactions (under 10 mm. Hg/min.) in which its greater accuracy was advantageous. Both gauges were encased in water-jackets to prevent them from being influenced by sudden changes in room temperature.

To prevent the condensation of steam by diffusion out of the reaction vessel during the course of a slow reaction, the entry tube (figure 2, *C*) was made of a 50 cm. coil of 2·5 mm. tubing, encased in a steam-jacket and maintained at 100° C. Since total pressure changes were to be measured as well as the reaction rates, the escape by condensation of steam formed in a reaction would affect the measurements. Typical curves of experiments made with and without this heater coil are shown in figure 3. Actually this diffusion of steam from the reaction vessel, with consequent condensation, only affects the pressure after several minutes. In the more rapid reactions, consequently, the maximum rate is achieved and passed before this diffusion can take effect. However, in the time taken to measure the total pressure change and in the slower reactions, say those requiring over 5 min. to reach maximum rate, this heater coil is of considerable importance

All the reaction vessels used in these tests were of Pyrex and were patterned to a standard length (approx. 20 cm). This length was chosen since, when centrally placed in the electric furnace, the vessel was exposed to a temperature constant within one degree over its entire length. The various vessels ranged from 3 to 40 mm. in diameter, the smaller vessels being made of a number of segments of the specified diameter attached in parallel so as, by avoiding great differences in volume, to obviate large correction factors for the volume of the gauge and its connecting tubing in computing the reaction velocity. A list of the vessels with their specifications follows (table 1).

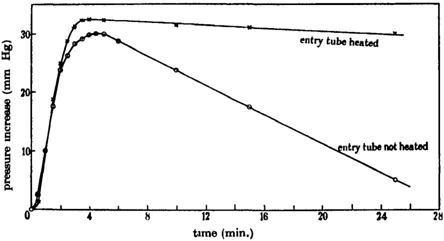


Figure 3. The effectiveness of heated entry tube as a preventive of steam condensation. $CH_4 + O_2$ at 530° C; initial pressure = 200 mm., vessel diameter = 29·0 mm.

Table 1

Internal diameter in mm.	No. of segments	Volume in c.c.
2 8	49	86·7
5·0	19	81·5
7·0	5	41 5
11 1	4	77·3
16 7	2	90·1
22·3	1	77 0
29 0	1	162 5
32 5	1	153 5
38 7	1	260
44 0	1	307

B. *Preparation of gases* (*a*) *Oxygen* was prepared by electrolysis of a 10 % solution of sodium hydroxide saturated with barium hydroxide, and purified from ozone and hydrogen by passage over heated (*ca.* 450° C) platinized asbestos (*b*) *Methane* was prepared by the reduction of methyl iodide, with methyl alcohol and a copper-zinc couple, mainly as described by Vanino (1914). Steps of purification were adopted principally from the method used by Norrish and Wallace (1934). (*c*) *Ethane* was made by electrolysis of a 20 % aqueous solution of potassium acetate, using platinum electrodes (Department of Scientific and Industrial Research, Fuel Research, 1931). (*d*) *Ethylene* was made by the dehydration of approximately 98 % ethyl alcohol in heated (220° C) phosphoric acid. The phosphoric acid was preliminarily heated somewhat higher (*ca.* 240° C) to remove excess water, and at all times during the preparation the temperature was not permitted to fall below 220° C, thus preventing the formation of ether. (*e*) *Propylene* was prepared by a method proposed by Sakmin (1934), in which propyl alcohol is dropped on to fragments of unglazed clay plate heated to 350° C. (*f*) *Acetylene* and *propane* were obtained from commercial cylinders.

Preliminary purification by chemical means was always followed by double distillation in vacuo at liquid air temperatures. All the gases prepared for the following tests were dried before use by passage over phosphorous pentoxide and were found to be at least 99 % pure by analysis on a modified Bone and Wheeler apparatus (1908; D.S.I.R. 1931).

C. *Standard procedure.* Throughout the entire series of experiments certain methods of procedure were standardized, whatever mixture of gases might be in use. Because of the extreme sensitivity of surface to many influences, a vessel when once installed was maintained at the desired temperature (within 1° C) until tests were completed. Also, air was under no circumstances permitted in a vessel after it had been evacuated and brought up to temperature, for the duration of the tests concerned. Before use each reaction vessel but one was washed with concentrated nitric acid and rinsed several times with distilled water, the exception being the 49 section, 2·8 mm. inside-diameter vessel, which was exempted because of its fragility. It had, however, been constructed of thoroughly cleaned tubing, and before use with any gas was kept, at least overnight, at a temperature in excess of 400° C while under vacuum. This treatment, it was assumed, would rid the surface of most impurities.

A desired combination of gases, when admitted to the mixing vessel, was given time for complete diffusion to take place—at least 15 min.—to ensure an homogeneous oxygen-hydrocarbon mixture.

All reaction rates were corrected for the volume of the Bourdon gauge and connecting tubing.

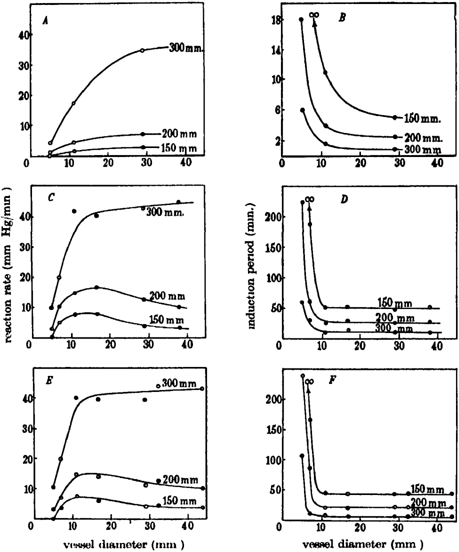
EXPERIMENTAL RESULTS

A. *Saturated hydrocarbons.* As a check on the operation of our own apparatus, as well as a measure of the reproducibility of surface experiments, part of the work of Norrish and Foord (1936) in the determination of the critical diameter for the slow oxidation of methane was repeated. The same conditions which had been used in the previous work were established, equimolecular mixtures of methane and oxygen at 530° C, and total pressures of 150, 200 and 300 mm. of mercury. The three vessels chosen for these tests were of 5·0, 11·1 and 29·0 mm. internal diameter and were treated before installation as described under paragraph C, above. The reaction rates obtained for the methane oxidation, when plotted against vessel diameter (figure 4, *A*), show a good reproduction of the previous work especially in the 5 mm. vessel, although appearing somewhat higher with the larger vessel diameters. This faster rate, however, may be attributable to any slight variation in surface conditions.

The oxidations of ethane and propane were undertaken in the following systems. $C_2H_6 + 2O_2$ and $C_3H_8 + 3O_2$, at 485° and 430° C, respectively, total pressures of 150, 200 and 300 mm. were also used, as for methane.

The reaction velocity fluctuated considerably in the course of the first few tests in any vessel, but became consistently steadier in each succeeding test. When its variation from the mean had reduced to about 8 %, the average of the final few values (3 or 4) was plotted against vessel diameter as illustrated by figure 4, *A*, *C*, *E*. The corresponding induction periods which were measured to the point of inflexion of the pressure-time curves (see below and figure 5), did not show the same variation as the reaction velocity, being much more reproducible, increasing somewhat during the first three or four tests they finally assumed a value that was constant within 15 sec for any number of succeeding tests. The induction periods are given in table 2 and plotted against vessel diameter in figure 4, *B*, *D*, *F*. Each reaction rate and its corresponding induction period represent the results of from eight to fourteen individual tests. Greater reproducibility of the velocity was obtained at the lower pressures and with smaller diameters; in general, the smaller the rate, the steadier were the values obtained.

The pressure-time curves for the three oxidations, from which reaction rates were calculated, were of the usual S-shaped type; examples are shown in figure 5. In all cases the pressure after having reached its maximum remained quite constant for some minutes, only decreasing very slowly as steam escaped from the reaction vessel through the heated entry tube and condensed in the cold external parts of the connecting tubing; this drop in pressure, as determined in a number of tests, was of the order of 5 mm. in 20 min. Illustration of the effectiveness of the heated entry tube as a preventive of steam condensation has been given in figure 3. The curves obtained for methane were similar to those found by Hinshelwood and Fort (1930), while the ethane curves were similar to those of Bone and Hill (1930)



A, *B*, $CH_4 + O_2$ at 530° C
C, *D*, $C_2H_6 + 2O_2$ at 485° C
E, *F*, $C_3H_8 + 3O_2$ at 430° C

FIGURE 4. Variation of slow reaction velocity and induction period with vessel diameter in the oxidation of saturated hydrocarbons.

for the same system; no case was found of the pressure rising to a maximum and then decreasing quite rapidly towards a negative value, as shown in the work of Kowalsky, Sadownikow and Tschirkow (1932). The only curves of that nature are those in which the entry tube was not heated.

The plot of the reaction rate against diameter shows a maximum in the oxidations of ethane and propane at the two lower pressures, 150 and 200 mm., at a diameter of about 12–16 mm. These values were repeated several times with frequent cleaning of the vessels, and no indication of contamination of the surface was found Several new vessels were made as well, all of which gave the same maximum in velocity. At 300 mm. pressure the rate approached a constant value as diameter increased, showing no maximum. The induction periods, however, showed no corresponding minimum at the lower pressures, in all cases assuming a steady value when the diameter was sufficiently greater than the critical value.

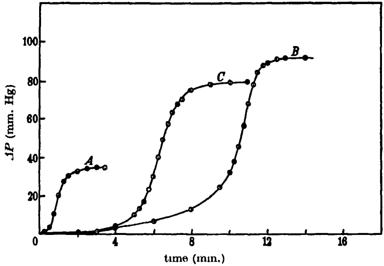


FIGURE 5 Representative reaction curves Increase of pressure with time in slow oxidation of saturated hydrocarbons. Initial pressure = 300 mm. Vessel diameter = 29·0 mm. *A*, $CH_4 + O_2$ at 530° C *B*, $C_2H_6 + 2O_2$ at 485° C. *C*, $C_3H_8 + 3O_2$ at 430° C.

Complete suppression of the reaction with all three gases was found in the 5·0 mm. vessel at 150 mm., no pressure change being observed in 20–24 hr., while the great reduction of rate and corresponding increase in induction period at higher pressures in the same vessel also demonstrate the existence of a critical diameter. Probably the most reliable means of calculating the net-branching factor ϕ is by use of the induction period, since its reproducibility throughout all tests was good, even when the reaction velocity tended to fluctuate considerably. From the relation $\phi\tau = \text{const.}$, relative values for ϕ have been computed as shown by ϕ/ϕ_0 in table 2, where the term ϕ_0 represents the value of ϕ corresponding to the average steady induction period found with large diameters at 300 mm. pressure. The values for propane are plotted against vessel diameter in figure 6; these curves, which are similar to those for methane and ethane (not plotted), give a very close approximation, especially at 300 mm. pressure, to the theoretical curve for ϕ shown in figure 1.

TABLE 2. VARIATION OF INDUCTION PERIOD (τ IN MINUTES) AND ϕ/ϕ_0 WITH VESSEL DIAMETER AT TOTAL PRESSURE P. ϕ_0 REPRESENTS AVERAGE VALUE FOR LARGE DIAMETERS AT 300 MM.

	Vessel diameter	$P = 300$ mm.		$P = 200$ mm		$P = 150$ mm.	
System	mm.	τ	ϕ/ϕ_0	τ	ϕ/ϕ_0	τ	ϕ/ϕ_0
$CH_4 + O_2$ at 530° C	29·0	0·8	1·00	2·5	0·32	5·1	0·16
	11·1	1·6	0·50	4·0	0·20	11·0	0·07
	5·0	6·0	0·13	18·0	0·04	∞	0·00
$C_2H_6 + 2O_2$ at 485° C	38·0	10·5	1·07	27·5	0·41	52·0	0·22
	29·0	11·0	1·03	26·5	0·43	48·0	0·23
	16·7	13·7	0·83	30·0	0·38	52·5	0·21
	11·1	10·0	1·13	25·0	0·45	51·3	0·22
	7·0	31·0	0·36	61·0	0·19	188·0	0·06
	5·0	60·0	0·19	223·0	0·05	∞	0·00
$C_3H_8 + 3O_2$ at 430° C	44·0	7·5	0·99	23·0	0·32	45·0	0·16
	32·5	7·5	0·99	23·5	0·31	44·0	0·17
	29·0	8·0	0·93	24·0	0·31	44·0	0·17
	16·7	6·5	1·14	21·5	0·34	44·0	0·17
	11·1	7·5	0·99	22·0	0·34	45·0	0·16
	7·0	11·0	0·67	86·5	0·09	166·0	0·04
	5·0	105·0	0·07	238·0	0·03	∞	0·00

Other methods of calculating ϕ have not been illustrated. However, computations based on the formula $\Delta p = Ne^{\phi t}$, plotting $\log(\Delta p)$ against time, give results in good qualitative agreement with the above method.

B. *Unsaturated hydrocarbons.* The three unsaturated hydrocarbons, acetylene, ethylene and propylene, were studied under the following conditions of admixture and temperature:

$$C_2H_2 + 1\tfrac{1}{2}O_2 \text{ at } 385° \text{ C},$$
$$C_2H_4 + 2O_2 \text{ at } 450° \text{ C},$$
$$C_3H_6 + 3O_2 \text{ at } 360° \text{ C}.$$

Total pressures of 150, 200 and 300 mm. were again used in these reactions.

Vessels larger than 29·0 mm. diameter were not used in the acetylene oxidation due to a tendency of the gases to explode violently in the mixing vessel. The explosions were unpredictable; several mixtures were detonated in the mixing vessel while using the 29·0 mm. reaction vessel, taking place at the moment when the tap between the two vessels was opened. Once a gas mixture had gained entrance to the reaction vessel without explosion, then the slow reaction proceeded normally. No explosions occurred when using vessels of less than 29·0 mm diameter.

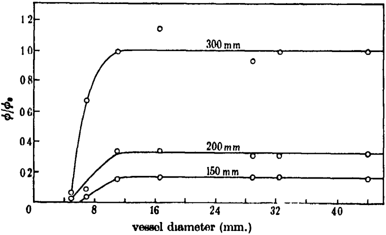
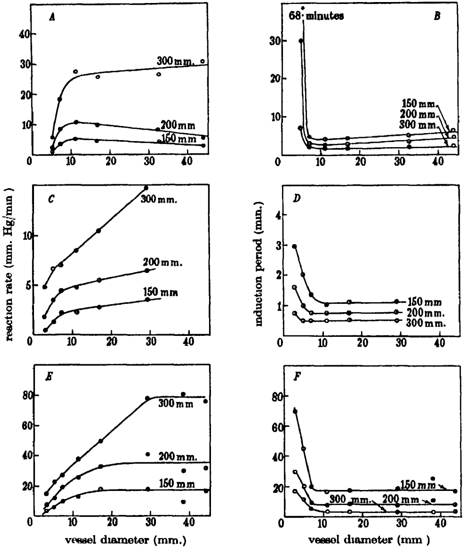


FIGURE 6. Variation of ϕ with vessel diameter in oxidation of propane. $C_3H_8 + 3O_2$ at 430° C.

Fluctuation of the reaction velocities in successive tests in the same vessel was not as great with acetylene and propylene as with ethylene, which acted very much the same as the saturated hydrocarbons in this respect. The first two gases gave reaction rates constant within about 5% variation from the mean after a vessel had been conditioned by a few preliminary tests. As a result, for each rate and its corresponding induction period, only six or seven tests in any vessel were required before the results were considered satisfactory; ethylene required ten or twelve tests to establish each point. Reaction rates and induction periods are plotted against vessel diameter as shown in figure 7, *A*–*F*.

Ethylene exhibits a maximum in the reaction velocity with pressures of 150 and 200 mm. at diameters of 12–16 mm. (figure 7, *A*), similarly to ethane and propane, tests with diameters from 12 to 40 mm. were repeated, using new vessels, with the maximum still apparent as indicated in the figure. Contrary to the behaviour of ethane and propane, a slight corre-



A, *B*, $C_2H_4 + 2O_2$ at 450° C
C, *D*, $C_2H_2 + \frac{3}{2}O_2$ at 385° C
E, *F*, $C_3H_6 + 3O_2$ at 360° C

FIGURE 7. Variation of slow reaction velocity and induction period with vessel diameter in the oxidation of unsaturated hydrocarbons.

sponding minimum is observed in the curves of the induction periods for ethylene (figure 7, *B*).

Complete suppression of the reaction was not obtained for these oxidations, a very small measurable reaction being found even in a vessel of 2·8 mm. diameter at 150 mm. pressure. Great inhibition of the reaction velocity is apparent, however, in all cases. Extrapolation of the reaction rate-vessel diameter curves (figure 7 *A*, *C*, *E*) shows that complete suppression of all homogeneous reaction may be expected at or about a diameter of 2 mm. with acetylene and propylene, while with ethylene it would appear to take place at about 4 mm.

The reaction rates of the acetylene and propylene oxidations are shown to decrease almost linearly between diameters of 29 and 7 mm., yet a correspondingly linear increase is not found in the curves for the induction periods, which show almost the same asymptotic increase with small diameters as was found for the saturates. Although very short when compared to the other gases, the induction periods for acetylene show a definite increase when the diameter is sufficiently decreased; this is contrary to Spence's observation (1932) that the induction period decreased when the surface/volume ratio was increased. Spence, though, used a flow method the results of which cannot be compared strictly with the static method employed in these tests.

The pressure-time curves obtained for these three oxidation reactions generally conformed with the S-type and were found to follow the form, $\Delta p = Ne^{\phi t}$, quite closely during the initial stages of the reaction (cf. figure 9). The reaction velocity shows a maximum which is most pronounced with propylene, but is quite distinct also in acetylene and ethylene (figure 8).

Values for ϕ/ϕ_0 have been calculated as before from the induction periods, and the same decrease of ϕ at a limiting diameter is observed. No example is plotted, but the trend of ϕ will be observed by reference to table 3.

DISCUSSION

It has been shown that the six oxidations investigated follow the form $\omega = Ae^{\phi t}$ very closely during the initial stages of reaction, for the plot of $\log(\Delta p)$ against time shows a straight line up to the point of inflexion of the pressure-time curves; this has been illustrated in figure 9 by values taken from the typical curves of figures 5 and 8. The curves of figure 9 are similar to those obtained from all the reactions studied in this investigation. Further, by measurement of the reaction velocity and induction period,

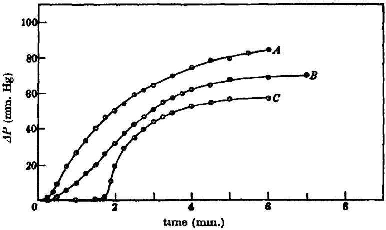


FIGURE 8. Representative reaction curves. Increase of pressure with time in slow oxidation of unsaturated hydrocarbons:

	Initial pressure mm. Hg	Vessel diameter mm.
A, $C_2H_2 + 1\frac{1}{2}O_2$ at 385° C ($\Delta p \times 5$)	300	11·1
B, $C_2H_4 + 2O_2$ at 450° C	300	44·0
C, $C_2H_6 + 3O_2$ at 360° C	200	44·0

TABLE 3. VARIATION OF INDUCTION PERIOD (τ IN MINUTES) AND ϕ/ϕ_0 WITH VESSEL DIAMETER AT TOTAL PRESSURE P. ϕ_0 REPRESENTS AVERAGE VALUE FOR LARGE DIAMETERS AT 300 MM.

	Vessel diameter	$P = 300$ mm.		$P = 200$ mm.		$P = 150$ mm.	
System	mm.	τ	ϕ/ϕ_0	τ	ϕ/ϕ_0	τ	ϕ/ϕ_0
$C_2H_4 + 2O_2$ at 450° C	44·0	2·2	0·84	4·4	0·42	6·0	0·31
	32·5	1·8	1·02	3·2	0·57	4·8	0·38
	16·7	1·6	1·15	2·6	0·71	4·2	0·44
	11·1	1·6	1·15	2·4	0·76	4·0	0·46
	7·0	2·0	0·92	2·8	0·66	4·6	0·40
	5·0	7·0	0·26	30·0	0·06	68·0	0·03
$C_2H_2 + 1\frac{1}{2}O_2$ at 385° C	29·0	0·5	1·00	0·75	0·67	1·1	0·45
	16·7	0·5	1·00	0·75	0·67	1·1	0·45
	11·1	0·5	1·00	0·75	0·67	1·0	0·50
	7·0	0·5	1·00	0·75	0·67	1·3	0·38
	5·0	0·5	1·00	1·00	0·50	2·0	0·25
	2·8	0·75	0·67	1·60	0·31	3·0	0·17
$C_2H_6 + 3O_2$ at 360° C	44·0	3·3	0·91	8·0	0·38	17·0	0·18
	38·0	2·9	1·03	11·0	0·27	25·5	0·12
	29·0	3·0	1·00	7·5	0·40	19·0	0·16
	16·7	3·0	1·00	8·5	0·35	17·5	0·17
	11·1	3·5	0·86	7·5	0·40	17·0	0·18
	7·0	5·5	0·55	9·5	0·32	20·0	0·16
	5·0	11·5	0·26	20·0	0·16	45·0	0·07
	2·8	17·0	0·18	30·0	0·10	70·0	0·04

the net-branching factor ϕ has been found to decrease to zero or negative values when the vessel diameter is reduced to a critical value. These results are in accordance with the theory of degenerate branching.

Chain reactions of the rapidly developing type, such as the oxidations of hydrogen and carbon monoxide, clearly demonstrate the operation of surface and volume deactivation. Upper and lower pressure limits of ignition define the extent to which these two types of deactivation may individually suppress the reaction despite all other factors leading to explosion. The two pressure limits of these reactions, which outline the explosion peninsula in pressure-temperature curves are shown to be dependent on the impedance of the gas molecules to the diffusion of chain centres. At the lower limit the pressure is sufficiently reduced to permit the chain centres access to the walls of the containing vessel where they are destroyed; as the pressure is increased, the diffusion of centres to the walls is impeded and the consequent increase in the concentration of centres results in almost instant inflammation; with further increase of pressure, a point is reached at which the destruction of active centres is promoted by gas-phase deactivation alone, and again explosive reaction is prevented.

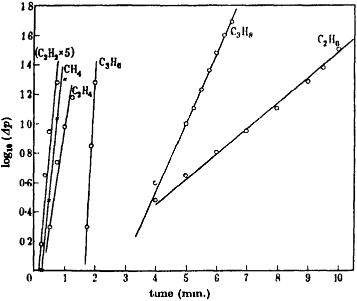


FIGURE 9. Representative curves of figures 5 and 8 with log ordinates.

With respect to upper and lower pressure limits of ignition, the oxidation reactions of the hydrocarbons do not show the same clear distinction of the two types of deactivation as do the rapidly developing chain reactions; although some evidence of a rudimentary separation of upper and lower limits is shown in ignition curves for methane-oxygen mixtures when the concentration of methane is between 2·5 and 37 % (Neumann and Serbinoff 1932), it would appear that the volume and surface deactivation are in general simultaneously operative factors, and the consequent differentiation into upper and lower limits is not pronounced with hydrocarbons. Reactions of the degenerate branching type may be expected to lack such differentiation at pressures below one atmosphere because the chain propagation, being effected both by way of atoms or radicals on the one hand, and more stable compounds such as aldehydes on the other, may be simultaneously exposed to control by influences operating both in the volume and at the surface.

In the present work it has been possible to demonstrate clearly the operation of both surface and volume deactivation in degenerate branching reactions as exemplified by the oxidations of hydrocarbons. Above a certain diameter of reaction vessel the reaction rate was found to attain a constant value independent of surface and limited by volume deactivation alone. However, when the surface/volume ratio is sufficiently increased, all other factors concerning chain branching become negligible by comparison, and the reaction is suppressed. As illustrated by equation (8),

$$\phi = f - g - \frac{\pi^2 D}{d^2}, \tag{8}$$

ϕ can be defined in terms of f, the secondary generation of centres from the oxidation of intermediate products, the volume deactivation g, and the surface deactivation S, expressed in terms of vessel diameter d and the gas diffusion coefficient D. When d is sufficiently large the surface deactivation becomes negligible, and the reaction velocity and induction period are constant, being dependent entirely on the factors controlling a homogeneous system (f and g) As d decreases $\pi^2 D/d^2$ becomes increasingly prominent until, at the critical diameter, ϕ becomes equal to or less than zero. Thus, when conditions of temperature, pressure, and gas composition are constant, equation (8) reduces to

$$\phi = a - \frac{b}{d^2}, \tag{9}$$

where a and b are constants and, through the limitation of ϕ by surface deactivation alone, the onset of the reaction is delayed infinitely. This has been demonstrated in the curves of figures 4 and 7.

The oxidation reactions of the hydrocarbons appear to be principally homogeneous, with little evidence of heterogeneity in the main reaction or of any residual surface reaction. The sharp decrease of ϕ near the critical diameter is approximately proportional to the square of diameter, which is in accordance with the initiation of chain centres from intermediate products in the volume and their destruction at the walls. In a heterogeneous reaction chain centres may be initiated both at the walls and in the volume. Hence, an increase in the surface/volume ratio will cause an increase in the number of centres created at the walls, and will counteract to some extent those being destroyed. Thus, the rate will vary with diameter according to some power less than the second, and under certain conditions will approximate a linear relationship. With acetylene and propylene the reaction rate decreases almost linearly between diameters of 29 and 7 mm. (figure 7) and might appear to support a heterogeneous reaction, but this linear dependency is not shown by ϕ when calculated from the induction periods; ϕ is seen to decrease very sharply near the limiting diameter as in the other four reactions.

The maxima found in the curves relating the reaction rate and vessel diameter for the ethane, propane and ethylene oxidations at 150 and 200 mm. pressure in vessels of 12–16 mm. diameter are probably due to some selective effect of surface. The maxima were not found when the pressure was increased to 300 mm., thus indicating that at lower pressures, when diffusion to and from the walls is less impeded, a certain heterogeneity of these reactions may be manifested. For example, in the oxidations of the higher hydrocarbons more complex intermediates can be assumed, which may be heterogeneous in their own oxidation reactions. The oxidation of formaldehyde, which is considered as an intermediate in the oxidation of methane, has been shown to be inhibited by an increase of surface (Askey 1930), but the rate of the acetaldehyde oxidation, studied by Hatcher, Steacie and Howland (1932) and by Pease (1933), was found to increase with increased surface/volume ratios under certain conditions. If the diameter is reduced in the oxidation of the higher hydrocarbons at low pressures, a critical condition may arise in which a heterogeneous reaction of an intermediate such as acetaldehyde is superimposed to a noticeable extent upon the main homogeneous reaction, thereby producing a maximum. At larger diameters the surface/volume ratio would be so reduced that this surface reaction would be negligible, while at very small diameters the limitation of the production of intermediate would reduce this secondary reaction as well.

There is the possibility of a small residual surface reaction being present in the oxidation of the unsaturated hydrocarbons, since complete inhibition of these reactions was not obtained with the smallest diameter, 2·8 mm., but the shape of the curves (figure 7) indicates that all observable reaction would cease with a further slight reduction in diameter.

REFERENCES

Askey 1930 *J. Amer. Chem. Soc.* **52**, 974.
Bone, Haffner and Rance 1933 *Proc. Roy. Soc.* A, **143**, 16.
Bone and Hill 1930 *Proc. Roy. Soc.* A, **129**, 434.
Bone and Wheeler 1908 *J. Soc. Chem. Ind., Lond.*, **27**, 10.
Department of Scientific and Industrial Research 1931 Fuel Research, Technical Paper no. 33.
Foord, S. G. 1934 *J. Sci. Instrum.* **11**, 126.
Foord and Norrish 1935 *Proc. Roy. Soc.* A, **152**, 196.
Hatcher, Steacie and Howland 1932 *Canad. J. Res.* **1**, 149.
Hinshelwood and Fort 1930 *Proc. Roy. Soc.* A, **129**, 284.
Kowalsky, Sadownikow and Tschirkow 1932 *Phys. Z. Sowjet.* **1**, 451.
Neumann and Serbinoff 1932 *Phys. Z. Sowjet.* **1**, 536.
Norrish 1935 *Proc. Roy. Soc.* A, **150**, 36.
Norrish and Foord 1936 *Proc. Roy. Soc.* A, **157**, 503.
Norrish and Wallace 1934 *Proc. Roy. Soc.* A, **145**, 307.
Pease, R. N. 1933 *J. Amer. Chem. Soc.* **55**, 2753.
Prettre, M. 1938 *C.R. Acad. Sci., Paris*, **207**, 532.
Sadownikow, P. 1937 *Acta phys.-chim. U.R.S.S.* **6**, 419.
Sakmin, P. K. 1934 *Ber. Dtsch. Chem. Ges.* **67** B, 392.
Semenoff, N. 1930 *Z. phys. Chem.* **11** B, 464.
— 1932 *Phys. Z. Sowjet.* **1**, 546.
— 1935 *Chemical kinetics and chain reactions.* Oxford.
Shantarovitch, P. 1937 *Acta Phys.-chim. U.R.S.S.* **7**, 417.
Spence 1932 *J. Chem. Soc.* p. 686.
Vanino 1914 *Präparative Chemie*, **2**, 1.

On the turbulence of a tidal current

By J. Proudman, F.R.S.

(*Received* 12 *June* 1940)

1. Introduction

This is a theoretical investigation of the dynamics of a tidal current when turbulence is of significance.

We take the depth of water to be uniform, and we suppose that the external body forces, the mean velocity of the current, and the statistics of turbulence are uniform over all horizontal planes, though they may vary between the sea surface and the sea bottom. In a natural tidal current there will usually be mean horizontal pressure gradients due to mean surface gradients. In order to maintain a horizontal mean sea surface, we shall suppose these pressure gradients to be replaced by external body forces.

In §§ 3–5 we take into consideration both the rotation of the earth and the oscillatory nature of a tidal current. We examine the forces of internal friction, the rates of doing work, and the energy relations, in terms of the statistics of the turbulence. The results are particular cases of those due to Osborne Reynolds (1894) and in fact the forces of internal friction consist largely of the Reynolds shearing stresses. In these sections no new principles are introduced, though most of the particular formulae of §§ 4, 5 do not appear to have been previously put on record.

In § 6 we emphasize that 'isotropic turbulence', which has recently been studied by G. I. Taylor (1935), T. de Karman and L. Howarth (1937), involves zero values of the Reynolds shearing stresses, and is therefore incapable of accounting for the quantities which are of interest in the present paper.

In §§ 7–11 we study certain particular types of disturbance. As these disturbances are specified by fairly concise formulae, they hardly correspond to the common idea of turbulence, but it is of interest to examine their contributions to the relationships investigated in §§ 3–5. As the fluctuations of velocity which constitute turbulence may be assumed to have zero mean values over periods which are very small compared with the tidal period, in §§ 7–11 we neglect both the tidal oscillation and the earth's rotation. As we are then studying a non-accelerated mean motion under external forces, there must be a balancing force due to friction at the sea bottom. This requires the non-vanishing of the forces of internal friction across any horizontal plane below the surface, but as the general expressions for the Reynolds stresses do not contain the coefficient of viscosity, it is conceivable that viscosity may be neglected except in the immediate neighbourhood of the sea bottom. Now in order to secure the desired constancy and uniformity in the statistics of the disturbance, we assume the time and the horizontal co-ordinates to enter only through trigonometrical functions. It then appears that it is only when viscosity is taken into account that the Reynolds shearing stresses do not vanish.

The disturbances we study consist partly of travelling vortices, and are similar to the oscillations of the usual theory of waves, except that the disturbance of the sea surface is very small. They are maintained against viscosity by a transference of energy, partly from the turbulent motion near the sea bottom, and partly from the local mean motion. Of course the turbulent energy near the sea bottom comes itself from that of the mean motion, but we do not investigate this particular transfer; and of course the mean motion in turn takes energy from the external body forces.

The investigations of §§ 6–11 may be regarded as a series of attempts to discover a type of disturbance which will produce non-zero values of the Reynolds shearing stresses. These attempts involve successively isotropic turbulence, motion without viscosity, and motion with viscosity, and only the last attempt is successful. Even then, the oscillations of real period which we study cannot extend right down to a perfectly plane sea bottom, but in any case such a sea bottom does not exist.

Apart from the conditions at the sea surface and sea bottom, the mathematics of §§ 7–11 is similar to that on the stability of laminar motion, on which there is a large literature.

There is not much in the paper that is specifically tidal; but the tides of shallow seas do appear to provide the best example in nature of the horizontal uniformity, the freedom from variations of density, and the significance of friction which are postulated throughout.

2. NOTATION AND FUNDAMENTAL EQUATIONS

We shall denote by:

ω	the vertical component of the earth's angular velocity,
g	the acceleration of gravity,
ρ	the density of the water,
μ	the coefficient of viscosity,
ν	the kinematic coefficient of viscosity, μ/ρ,
t	the time,

x, y, z	Cartesian co-ordinates, with the origin O in the mean sea surface and Oz vertically upwards,
U, V	the components of the mean current,
X, Y	the components of the external body force per mass,
P	the mean pressure,
N	the eddy-viscosity function, when the mean motion is confined to one vertical plane,
u, v, w	the components of the 'turbulent velocity',
q	the speed of the turbulent motion, such that $q^2 = u^2+v^2+w^2$,
χ	the vertical component of the 'turbulent vorticity',
p	the 'turbulent pressure',
ζ	the 'turbulent elevation' of the sea surface,
$\mu\phi$	the rate of dissipation of 'turbulent energy' by viscosity.

The quantities from ω to ν we take to be constants; the quantities from U, V to N we take to be functions of t and z; the quantities from u, v, w to p we take to be functions of t and x, y, z; while ζ we take to be a function of t and x, y.

We shall use the symbol [ ] to denote a mean value which eliminates any one of the quantities from u, v, w to ζ above, but which makes no appreciable difference in any of the quantities from U, V to N above.

From the equation of continuity we have

$$\frac{\partial U}{\partial x}+\frac{\partial V}{\partial y}=0, \quad \frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0. \tag{2·1}$$

3. Equations of motion

The equations of mean motion are

$$\left.\begin{aligned}\frac{\partial U}{\partial t}-2\omega V+\frac{\partial}{\partial z}[uw] &= \nu\frac{\partial^2 U}{\partial z^2}+X,\\ \frac{\partial V}{\partial t}+2\omega U+\frac{\partial}{\partial z}[vw] &= \nu\frac{\partial^2 V}{\partial z^2}+Y,\\ \frac{\partial}{\partial z}[w^2] &= -\frac{1}{\rho}\frac{\partial P}{\partial z}-g.\end{aligned}\right\} \tag{3·1}$$

From the third equation of (3·1) we deduce

$$\frac{P}{\rho}=-gz-[w^2]+f(t), \tag{3·2}$$

where $f(t)$ is a function of t only.

The quantities $\quad -\rho[uw], \quad -\rho[vw], \quad -\rho[w^2]$

are constituents of the Reynolds stress system; they govern the interaction between the mean motion and the turbulent motion. If U, V, X, Y are harmonic functions of t, then $[uw]$ and $[vw]$ will also be harmonic functions of t with the same period.

The equations of turbulent motion are

$$\left.\begin{aligned}
\frac{\partial u}{\partial t}+(U+u)\frac{\partial u}{\partial x}+(V+v)\frac{\partial u}{\partial y}+w\frac{\partial}{\partial z}(U+u)-\frac{\partial}{\partial z}[uw]-2\omega v &= -\frac{1}{\rho}\frac{\partial p}{\partial x}+\nu\nabla^2 u,\\
\frac{\partial v}{\partial t}+(U+u)\frac{\partial v}{\partial x}+(V+v)\frac{\partial v}{\partial y}+w\frac{\partial}{\partial z}(V+v)-\frac{\partial}{\partial z}[vw]+2\omega u &= -\frac{1}{\rho}\frac{\partial p}{\partial y}+\nu\nabla^2 v,\\
\frac{\partial w}{\partial t}+(U+u)\frac{\partial w}{\partial x}+(V+v)\frac{\partial w}{\partial y}+w\frac{\partial w}{\partial z}-\frac{\partial}{\partial z}[w^2] &= -\frac{1}{\rho}\frac{\partial p}{\partial z}+\nu\nabla^2 w.
\end{aligned}\right\} \tag{3·3}$$

In §§ 7–11 we shall take

$$\omega = 0, \quad V = 0, \quad Y = 0, \quad \partial U/\partial t = 0, \quad \partial X/\partial t = 0,$$

and then the first two of equations (3·1) become

$$\frac{d}{dz}[uw] = \nu\frac{d^2U}{dz^2}+X, \quad \frac{d}{dz}[vw] = 0. \tag{3·4}$$

Integrating these, we obtain

$$[uw]-\nu\frac{dU}{dz} = \int_z^0 X\,dz, \quad [vw] = 0, \tag{3·41}$$

on assuming the sea surface to be free from stress. In the tides X is independent of z, so that, where $\nu dU/dz$ is negligible, $[uw]$ is proportional to the depth below the surface.

We shall define the eddy-viscosity function N by the equation

$$N = \nu-\frac{[uw]}{dU/dz}, \tag{3·5}$$

so that the mean value of the total horizontal shearing stress is $\rho N\,dU/dz$ per area.

4. RATES OF DOING WORK

We shall first consider the mean rates of doing work across a horizontal plane by the stresses in the plane as they act on the water below the plane.

The rate of doing work, per area, by the mean stresses is

$$U\left(\mu\frac{\partial U}{\partial z}-\rho[uw]\right)+V\left(\mu\frac{\partial V}{\partial z}-\rho[vw]\right). \tag{4·1}$$

The mean rate of doing work, per area, by the turbulent stresses is

$$-[wp]+\mu\frac{\partial}{\partial z}[\tfrac{1}{2}q^2+w^2]. \tag{4·2}$$

We shall next consider the mean rates of doing work on an element of water.

The rate of doing work, per volume, by the resultant mean stresses on an element is

$$U\left(\mu\frac{\partial^2 U}{\partial z^2}-\rho\frac{\partial}{\partial z}[uw]\right)+V\left(\mu\frac{\partial^2 V}{\partial z^2}-\rho[vw]\right). \tag{4·3}$$

The mean rate of doing work, per volume, by the resultant turbulent stresses on an element is

$$-\frac{\partial}{\partial z}[wp]+\mu\frac{\partial^2}{\partial z^2}[\tfrac{1}{2}q^2+w^2]-\mu[\phi], \tag{4·4}$$

where

$$\phi=2\left(\frac{\partial u}{\partial x}\right)^2+2\left(\frac{\partial v}{\partial y}\right)^2+2\left(\frac{\partial w}{\partial z}\right)^2+\left(\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y}\right)^2+\left(\frac{\partial w}{\partial x}+\frac{\partial u}{\partial z}\right)^2+\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)^2.$$

The rate, per volume, at which energy is lost to the mean motion is

$$\frac{\partial}{\partial z}\left\{U\left(\mu\frac{\partial U}{\partial z}-\rho[uw]\right)+V\left(\mu\frac{\partial V}{\partial z}-\rho[vw]\right)\right\}$$
$$-\left\{U\left(\mu\frac{\partial^2 U}{\partial z^2}-\rho\frac{\partial}{\partial z}[uw]\right)+V\left(\mu\frac{\partial^2 V}{\partial z^2}-\rho\frac{\partial[vw]}{\partial z}\right)\right\},$$

and this is equal to

$$\mu\left\{\left(\frac{\partial U}{\partial z}\right)^2+\left(\frac{\partial V}{\partial z}\right)^2\right\}-\rho\left\{\frac{\partial U}{\partial z}[uw]+\frac{\partial V}{\partial z}[vw]\right\}. \tag{4·5}$$

The first term of (4·5) is the rate of dissipation, by viscosity, of the energy of mean motion, and is usually negligible except in the immediate neighbourhood of the sea bottom. The second term of (4·5) is the rate at which energy is transferred from mean motion to turbulent motion.

The mean rate, per volume, at which energy is lost to turbulence is

$$\mu[\phi],$$

which is the mean rate of dissipation, by viscosity, of the energy of turbulence.

5. ENERGY EQUATIONS

The energy equation for turbulence is

$$\frac{\partial}{\partial t}[\tfrac{1}{2}q^2]+\frac{\partial}{\partial z}[w\,.\,\tfrac{1}{2}q^2]+[uw]\frac{\partial U}{\partial z}+[vw]\frac{\partial V}{\partial z}$$
$$=-\frac{1}{\rho}\frac{\partial}{\partial z}[wp]+\nu\frac{\partial^2}{\partial z^2}[\tfrac{1}{2}q^2+w^2]-\nu[\phi], \qquad (5\cdot1)$$

the terms on the right-hand side being the same as (4·4) divided by ρ.

Now integrate the equation (5·1) between the levels $z=z_1$ and $z=z_2$; we obtain

$$\int_{z_1}^{z_2}\left\{\frac{\partial}{\partial t}(\tfrac{1}{2}q^2)+[uw]\frac{\partial U}{\partial z}+[vw]\frac{\partial V}{\partial z}+\nu[\phi]\right\}dz$$
$$=\left|-[w\,.\,\tfrac{1}{2}q^2]-\frac{1}{\rho}[wp]+\nu\frac{\partial}{\partial z}[\tfrac{1}{2}q^2+w^2]\right|_{z_1}^{z_2}, \qquad (5\cdot2)$$

where $\Big|\Big|_{z_1}^{z_2}$ denotes the excess of the value of the function at $z=z_1$ over that at $z=z_2$. On using the equations (3·1) we obtain from (5·2)

$$\int_{z_1}^{z_2}\left\{\frac{\partial}{\partial t}\tfrac{1}{2}(U^2+V^2+q^2)-(XU+YV)+\nu\left\{\left(\frac{\partial U}{\partial z}\right)^2+\left(\frac{\partial V}{\partial z}\right)^2+[\phi]\right\}\right\}dz$$
$$=\left|-[w\,.\,\tfrac{1}{2}q^2]+U\left(\nu\frac{\partial U}{\partial z}-[uw]\right)+V\left(\nu\frac{\partial V}{\partial z}-[vw]\right)-\frac{1}{\rho}[wp]+\nu\frac{\partial}{\partial z}[\tfrac{1}{2}q^2+w^2]\right|_{z_1}^{z_2}. \qquad (5\cdot3)$$

The term $[w\,.\,\tfrac{1}{2}q^2]$ represents the vertical convection of the energy of turbulence per density; the interpretations of the other terms on the right-hand sides of (5·2) and (5·3) are given by (4·1) and (4·2).

If we take $z=z_2$ to correspond to the mean surface of the sea, and suppose this surface to be free from stress, then the contributions from this level to the right-hand sides of (5·2) and (5·3) will vanish.

If we could also regard the sea bottom as a perfect horizontal plane, then every term on the right-hand sides of (5·2) and (5·3) would separately vanish. But a more practical formula will be obtained by taking $z=z_1$ to correspond to a level a few centimetres above the bottom.

6. Isotropic turbulence

At any point and time, let u' be the component of turbulent velocity in a direction whose cosines are l, m, n. Then 'isotropic turbulence' is defined by the condition that $[u'^2]$ shall be independent of l, m, n.

Now
$$u' = lu + mv + nw,$$
so that
$$[u'^2] = l^2[u^2] + m^2[v^2] + n^2[w^2] + 2mn[vw] + 2nl[wu] + 2lm[uv],$$
and this can only be independent of l, m, n when
$$[u^2] = [v^2] = [w^2],$$
$$[vw] = [wu] = [uv] = 0.$$

We thus see that the shearing constituents of the Reynolds stress system are zero, so that isotropic turbulence contributes nothing to internal friction. In particular, we see from (4·5) that in our problem no energy can be transferred from mean motion to turbulent motion.

7. Disturbance of constant shearing current

If we neglect the rotation of the earth and the time variation of the mean current, and suppose that the direction of this current is the same at all depths, we may take
$$\omega = 0, \quad V = 0, \quad \frac{\partial U}{\partial t} = 0.$$

The equations (3·3) may then be written
$$\left.\begin{aligned}
&\frac{\partial u}{\partial t} + U\frac{\partial u}{\partial x} + w\frac{dU}{dz} - \nu\nabla^2 u + \frac{1}{\rho}\frac{\partial p}{\partial x}\\
&\qquad\qquad = -\left(u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z}\right) + \frac{d}{dz}[uw],\\
&\frac{\partial v}{\partial t} + U\frac{\partial v}{\partial x} - \nu\nabla^2 v + \frac{1}{\rho}\frac{\partial p}{\partial y} = -\left(u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z}\right),\\
&\frac{\partial w}{\partial t} + U\frac{\partial w}{\partial x} - \nu\nabla^2 w + \frac{1}{\rho}\frac{\partial p}{\partial z} = -\left(u\frac{\partial w}{\partial x} + v\frac{\partial w}{\partial y} + w\frac{\partial w}{\partial z}\right) + \frac{d}{dz}[w^2],
\end{aligned}\right\} \qquad (7·1)$$

and from these and the equation of continuity (2·1) we deduce

$$\left(\nu\nabla^2-\frac{\partial}{\partial t}-U\frac{\partial}{\partial x}\right)\nabla^2 w+\frac{d^2U}{dz^2}\frac{\partial w}{\partial x}=\nabla^2\left(u\frac{\partial w}{\partial x}+v\frac{\partial w}{\partial y}+w\frac{\partial w}{\partial z}\right)-\frac{\partial S}{\partial z}, \quad (7\cdot21)$$

$$\left(\nu\nabla^2-\frac{\partial}{\partial t}-U\frac{\partial}{\partial x}\right)\chi+\frac{dU}{dz}\frac{\partial w}{\partial y}$$
$$=u\frac{\partial \chi}{\partial x}+v\frac{\partial \chi}{\partial y}+w\frac{\partial \chi}{\partial z}+\frac{\partial w}{\partial x}\frac{\partial v}{\partial z}-\frac{\partial w}{\partial y}\frac{\partial u}{\partial z}-\frac{\partial w}{\partial z}\chi, \quad (7\cdot22)$$

$$\nabla_1^2\frac{p}{\rho}+\left(\nu\nabla^2-\frac{\partial}{\partial t}-U\frac{\partial}{\partial x}\right)\frac{\partial w}{\partial z}+\frac{dU}{dz}\frac{\partial w}{\partial x}=\left(u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}+w\frac{\partial}{\partial z}\right)\frac{\partial w}{\partial z}-S_1, \quad (7\cdot23)$$

where
$$\chi=\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}, \quad \nabla_1^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2},$$

$$S=\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial v}{\partial y}\right)^2+\left(\frac{\partial w}{\partial z}\right)^2+2\left(\frac{\partial w}{\partial y}\frac{\partial v}{\partial z}+\frac{\partial u}{\partial z}\frac{\partial w}{\partial x}+\frac{\partial v}{\partial x}\frac{\partial u}{\partial y}\right),$$

$$S_1=\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial v}{\partial y}\right)^2+\frac{\partial w}{\partial y}\frac{\partial v}{\partial z}+\frac{\partial u}{\partial z}\frac{\partial w}{\partial x}+2\frac{\partial v}{\partial x}\frac{\partial u}{\partial y}.$$

At the surface of the sea we have

$$\left.\begin{aligned}&P+p-2\mu\frac{\partial u_n}{\partial n}=\text{atmospheric pressure,}\\&\nu\left(\frac{\partial u_s}{\partial n}+\frac{\partial u_n}{\partial s}\right)=0,\end{aligned}\right\} \quad (7\cdot3)$$

where u_n, u_s denote the components of turbulent velocity along the normal and any tangent to the surface, while $\partial/\partial n$, $\partial/\partial s$ denote differentiation in these two directions. Assuming the atmospheric pressure to be uniform and unaffected by the disturbance in the water, we have, on using (3·2),

$$\frac{p}{\rho}-2\nu\frac{\partial u_n}{\partial n}=gz+[w^2]+\text{constant}$$

at $z=\zeta$. From this we deduce

$$\left(\frac{\partial}{\partial t}+U\frac{\partial}{\partial x}\right)\left(\frac{p}{\rho}-2\nu\frac{\partial u_n}{\partial n}\right)-gw$$
$$=-\left(u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}+w\frac{\partial}{\partial z}\right)\left(\frac{p}{\rho}-2\nu\frac{\partial u_n}{\partial n}\right)+w\frac{d}{dz}[w^2] \quad (7\cdot31)$$

and
$$g\zeta = \frac{p}{\rho} - 2\nu\frac{\partial u_n}{\partial n} - [w^2] + \text{constant} \tag{7·32}$$
at $z = \zeta$.

It is our intention to study disturbances which are periodic in t, x, y. If we were to take $u = v = w = 0$ at the sea bottom, the first order terms in the above equations would then be the same as for ordinary wave motion, and we should obtain oscillations damped by viscosity. But we wish to study oscillations which are maintained against viscosity by the external body forces acting through the mean motion.

From (3·41) we have
$$[uw] - \nu\frac{dU}{dz} = 0 \quad (z = 0),$$
and though observation indicates that dU/dz may not vanish at the sea surface, yet in all cases $[uw]$ must be small at the surface compared with values which it takes below the surface. Also, we do not wish to study surface waves any further than they may be necessarily involved.

We therefore take the following boundary conditions, in addition to (7·3) and (7·31),
$$w = 0 \quad (z = 0), \tag{7·41}$$
$$[uw] = T \quad (z = -h), \tag{7·42}$$
$$\frac{1}{\rho}[wp] - \nu\frac{d}{dz}[\tfrac{1}{2}q^2 + w^2] = W \quad (z = -h), \tag{7·43}$$
where T and W are prescribed constants.

The condition (7·41) will exclude surface waves to the first order. Of course, in the actual tides, surface waves will usually be present, but our object is to study disturbances which are as simple as the possibility of having non-zero forces of internal friction allows.

The level $z = -h$ is supposed to be above the sea bottom. The quantity ρT represents the Reynolds shearing stress by which the water below the level $z = -h$ acts on the water above that level. The quantity ρW represents the mean rate at which, through the turbulent stresses, the water below the level $z = -h$ does work on the water above that level.

8. PERIODIC DISTURBANCE, NEGLECTING VISCOSITY

When we neglect ν and the terms on the right-hand sides of (7·21), (7·22), (7·23) and (7·31), these equations possess a solution of the form

$$\left.\begin{aligned} u &= u_1 \sin(\sigma t+\kappa x)\cos ky,\\ v &= v_1 \cos(\sigma t+\kappa x)\sin ky,\\ w &= w_1 \cos(\sigma t+\kappa x)\cos ky,\\ \chi &= \chi_1 \sin(\sigma t+\kappa x)\sin ky,\\ p &= p_1 \sin(\sigma t+\kappa x)\cos ky,\\ \zeta &= \zeta_1 \sin(\sigma t+\kappa x)\cos ky,\end{aligned}\right\} \qquad (8\cdot1)$$

where σ, κ and k are constants, where u_1, v_1, w_1, χ_1 and p_1 are functions of z only, and where ζ_1 is a constant. But when we come to consider the right-hand sides of (7·21) and (7·22) we shall see that the disturbance is either irrotational or is restricted to the two dimensions of x and z.

We have from (2·1), (7·21), (7·22) and (7·23)

$$\left\{(\sigma+\kappa U)\left(\frac{d^2}{dz^2}-\kappa^2-k^2\right)-\kappa\frac{d^2U}{dz^2}\right\}w_1 = 0, \qquad (8\cdot21)$$

$$(\sigma+\kappa U)\chi_1 = -k\frac{dU}{dz}w_1, \qquad (8\cdot22)$$

$$\left.\begin{aligned} \kappa u_1+kv_1 &= -\frac{dw_1}{dz},\\ ku_1-\kappa v_1 &= \chi_1,\end{aligned}\right\} \qquad (8\cdot23)$$

$$(\kappa^2+k^2)\frac{p_1}{\rho} = (\sigma+\kappa U)\frac{dw_1}{dz}-\kappa\frac{dU}{dz}w_1. \qquad (8\cdot24)$$

To the first order, the equation (7·31) yields

$$\left(\frac{\partial}{\partial t}+U\frac{\partial}{\partial x}\right)\frac{p}{\rho} = gw \quad (z=0),$$

and then from this and (7·41) we have

$$w_1 = 0, \quad (\sigma+\kappa U)p_1 = 0 \quad (z=0). \qquad (8\cdot3)$$

The second condition of (8·3) shows that either $\sigma+\kappa U$ or p_1 must vanish at $z=0$, but we shall proceed to show that both these quantities must vanish at $z=0$.

Except where $\sigma+\kappa U$ already vanishes for a continuous range of values of z, w_1 is determined by a homogeneous differential equation of the second order, viz. (8·21), and the vanishing of both w_1 and dw_1/dz at $z=0$ would require the vanishing of w_1 everywhere. Hence, from the first condition of (8·3), dw_1/dz cannot vanish at $z=0$. From (8·24) and (8·3) we have

$$\frac{(\sigma+\kappa U)^2}{\kappa^2+k^2}\frac{dw_1}{dz}=(\sigma+\kappa U)\frac{p_1}{\rho}=0 \quad (z=0),$$

and hence $$\sigma+\kappa U=0 \quad (z=0). \tag{8·41}$$

From (8·24) and the first condition of (8·3) we then see that

$$p_1=0 \quad (z=0). \tag{8·42}$$

The equation (7·32) gives $g\zeta_1=\dfrac{p_1}{\rho} \quad (z=0),$

and hence $$\zeta_1=0. \tag{8·43}$$

On writing U_0 for the value of U when $z=0$, we see from (8·41) that t and x only enter through the form $\kappa(x-U_0t)$, so that the disturbance travels with the velocity of the surface current.

Equations (8·22), (8·23) and (8·24) determine u_1, v_1, p_1 in terms of w_1. The solution of (8·21) involves two constants of integration in w_1, one of these is required to satisfy the first equation of (8·3), while the other determines the magnitude of the whole disturbance.

If in (8·1) we had taken

$$u=\{u_1\sin(\sigma t+\kappa x)+u_1'\cos(\sigma t+\kappa x)\}\cos ky$$

with corresponding terms in v_1' and w_1', then u_1', v_1', w_1' would satisfy the same equations (8·21)–(8·43) as u_1, v_1, w_1. From the differential equation (8·21) and the surface condition (8·3), w_1' would only differ from w_1 by a constant factor, and it would then follow that

$$\frac{u_1'}{u_1}=\frac{v_1'}{v_1}=\frac{w_1'}{w_1}=\text{constant}.$$

The introduction of u_1', v_1', w_1' would thus give a phase change independent of z, and this would not constitute any essential generalization.

We next substitute from (8·1) into the right-hand sides of the equations of § 7; from (7·21) we obtain

$$\frac{1}{2}\left(w_1\frac{d^3w_1}{dz^3}-\frac{dw_1}{dz}\frac{d^2w_1}{dz^2}\right)\cos 2(\sigma t+\kappa x)\cos 2ky$$
$$+\kappa\left(-\tfrac{1}{2}\kappa u_1\frac{du_1}{dz}+\tfrac{1}{2}ku_1\frac{dv_1}{dz}+kv_1\frac{du_1}{dz}-\tfrac{1}{2}w_1\frac{d^2u_1}{dz^2}+2\kappa kv_1w_1\right)\cos 2(\sigma t+\kappa x)$$
$$+k\left(-\tfrac{1}{2}kv_1\frac{dv_1}{dz}+\tfrac{1}{2}\kappa v_1\frac{du_1}{dz}+\kappa u_1\frac{dv_1}{dz}-\tfrac{1}{2}w_1\frac{d^2v_1}{dz^2}+2\kappa ku_1w_1\right)\cos 2ky, \qquad (8{\cdot}51)$$

and from (7·22) we obtain

$$\frac{1}{2}\left(w_1\frac{d\chi_1}{dz}-\chi_1\frac{dw_1}{dz}\right)\sin 2(\sigma t+\kappa x)\sin 2ky. \qquad (8{\cdot}52)$$

When $\sigma+\kappa U$ vanishes for all values of z, the equations (8·21) and (8·22) place no restrictions on w_1 and χ_1, but the right-hand sides of (7·21) and (7·22) must then vanish. This requires

$$w_1\frac{d^3w_1}{dz^3}-\frac{dw_1}{dz}\frac{d^2w_1}{dz^2}=0, \quad w_1\frac{d\chi_1}{dz}-\chi_1\frac{dw_1}{dz}=0,$$

from which we deduce that

$$\frac{d^2w_1/dz^2}{w_1} \quad \text{and} \quad \frac{\chi_1}{w_1}$$

must both be constants, and then from the remaining terms of (8·51) we see, after some reduction, that either $k=0$, or the disturbance must be irrotational.

Now suppose that $\sigma+\kappa U$ does not vanish for a continuous range of values of z, and let us seek a second approximation of the same general form as (8·1). We require an expression for w which, when substituted in the left-hand side of (7·21), will produce (8·51). But this can only be found when the coefficient of $\cos 2ky$ in (8·51) is zero. On using the equations (8·21), (8·22) and (8·23) the coefficient of $-\kappa k^2\cos 2ky$ in (8·51) may be written as

$$\frac{dU/dz}{\sigma+\kappa U}\left\{w_1^2+\frac{\kappa^2-k^2}{(\kappa^2+k^2)^2}\left(\frac{dw_1}{dz}-\frac{dU/dz}{\sigma+\kappa U}w_1\right)^2\right\},$$

and this can only vanish for special constant values of

$$\frac{dU/dz}{\sigma+\kappa U}.$$

As $\sigma+\kappa U_0=0$, the only applicable case is that of a constant value of U and this has already been examined. But when $k=0$ no restriction need be placed on U.

For the two-dimensional case in which $k = 0$, a second approximation of the same general form as (8·1) exists.

On utilizing such relationships as

$$[\cos(\sigma t+\kappa x)\sin(\sigma t+\kappa x)] = [\cos(\sigma t+\kappa x)\cos 2(\sigma t+\kappa x)] = 0,$$

we see that

$$\left.\begin{aligned}[vw] = [wu] = [uv] = 0,\\ [w.\tfrac{1}{2}q^2] = 0, \quad [wp] = 0.\end{aligned}\right\} \tag{8·6}$$

If we start with the sum of a number of solutions of the type (8·1) and work out the second approximations, we see that the results (8·6) remain true.

We cannot, therefore, satisfy either of the conditions (7·42) and (7·43) when T and W are different from zero. Also, the rates of doing work (4·1), (4·2), (4·3), (4·4), (4·5) are all zero, and in the energy equations (5·1), (5·2), (5·3) every term is separately zero.

9. Examples of periodic disturbances, neglecting viscosity

Although the results of § 8 show that disturbances of the type there considered cannot account for the phenomenon of internal friction due to turbulence, yet it is of interest to consider certain specially simple cases. We shall give three examples of two-dimensional disturbances in each of which the expressions satisfy the differential equations to all orders, but the surface conditions only to the second order.

We shall take c to denote a constant speed fixing the magnitude of each disturbance.

(1) Take $U = U_0$, where U_0 is a constant, so that the whole disturbance is convected by the mean current. A possible disturbance is given by*

$$\frac{u}{c} = -\lambda \sin\kappa(x-U_0t)\cos\lambda\kappa z,$$

$$\frac{w}{c} = \cos\kappa(x-U_0t)\sin\lambda\kappa z,$$

$$\frac{p}{\rho c^2} = \tfrac{1}{4}\lambda^2\cos 2\kappa(x-U_0t),$$

$$\frac{g\zeta}{c^2} = \tfrac{1}{4}\lambda^2\cos 2\kappa(x-U_0t),$$

---

* Cf. C. L. Godske (1934).

for all values of the constants κ and λ. The disturbance may be regarded as a system of vortex cells.

(2) Now take
$$U = U_0 + \gamma z,$$
where U_0 and γ are constants. The equation (7·21) shows that the first-order disturbances will be irrotational. A possible disturbance is given by

$$\frac{u}{c} = -\sin\kappa(x-U_0t)\cosh\kappa z,$$

$$\frac{w}{c} = \cos\kappa(x-U_0t)\sinh\kappa z,$$

$$\frac{\kappa p}{\gamma\rho} = -\sin\kappa(x-U_0t)(\sinh\kappa z-\kappa z\cosh\kappa z)+\frac{\kappa c^2}{4\gamma}\cos 2\kappa(x-U_0t),$$

$$\frac{g\zeta}{c^2} = \tfrac{1}{4}\cos 2\kappa(x-U_0t),$$

for any value of the constant κ.

(3) Next take
$$U = U_0 + U_1\sin mz,$$
where U_0, U_1 and m are constants. A possible disturbance is given by

$$\frac{u}{c} = -\frac{\lambda}{\kappa}\sin\kappa(x-U_0t)\cosh\lambda z,$$

$$\frac{w}{c} = \cos\kappa(x-U_0t)\sinh\lambda z,$$

$$\frac{p}{\rho U_1 c} = \sin\kappa(x-U_0t)\left(\frac{\lambda}{\kappa}\sin mz\cosh\lambda z-\frac{m}{\kappa}\cos mz\sinh\lambda z\right) + \frac{1}{4}\frac{c}{U_1}\frac{\lambda^2}{\kappa^2}\cos 2\kappa(x-U_0t),$$

$$\frac{g\zeta}{c^2} = \frac{1}{4}\frac{\lambda^2}{\kappa^2}\cos 2\kappa(x-U_0t),$$

for all values of the constants κ and λ satisfying
$$\kappa^2 = m^2+\lambda^2.$$

This disturbance has a vorticity.

All three examples involve second-order disturbances of the sea surface.

10. PERIODIC DISTURBANCE, ALLOWING FOR VISCOSITY

Having so far failed to account for non-zero forces of internal friction, we now return to the general equations of § 7 and retain the coefficient of viscosity ν.

On neglecting terms on the right-hand sides of (7·21), (7·22), (7·23) and (7·31), these equations possess a solution of the form

$$\left.\begin{aligned} u &= u_1 \cos ky \\ v &= v_1 \sin ky \\ w &= w_1 \cos ky \\ \chi &= \chi_1 \sin ky \\ p &= p_1 \cos ky \\ \zeta &= \zeta_1 \cos ky \end{aligned}\right\} e^{i(\sigma t+\kappa x)}, \tag{10·1}$$

where σ, κ and k are real constants, where u_1, v_1, w_1, χ_1 and p_1 are functions of z only, and where ζ_1 is a constant. We thus work with complex harmonic motion, so that we shall have to interpret only the real parts. From the equations above mentioned we have

$$\left[\left\{\nu\left(\frac{d^2}{dz^2}-\kappa^2-k^2\right)-i(\sigma+\kappa U)\right\}\left\{\frac{d^2}{dz^2}-\kappa^2-k^2\right\}+i\kappa\frac{d^2U}{dz^2}\right]w_1 = 0, \tag{10·21}$$

$$\left\{\nu\left(\frac{d^2}{dz^2}-\kappa^2-k^2\right)-i(\sigma+\kappa U)\right\}\chi_1 = k\frac{dU}{dz}w_1, \tag{10·22}$$

$$\left.\begin{aligned} i\kappa u_1+kv_1 &= -\frac{dw_1}{dz}, \\ ku_1+i\kappa v_1 &= \chi_1, \end{aligned}\right\} \tag{10·23}$$

$$(\kappa^2+k^2)\frac{p_1}{\rho} = \left\{\nu\left(\frac{d^2}{dz^2}-\kappa^2-k^2\right)-i(\sigma+\kappa U)\right\}\frac{dw_1}{dz}+i\kappa\frac{dU}{dz}w_1, \tag{10·24}$$

and from (7·3) and (7·41),

$$w_1 = 0, \quad \frac{du_1}{dz} = 0, \quad \frac{dv_1}{dz} = 0 \quad (z = 0), \tag{10·31}$$

while from (7·31) and (7·32)

$$(\sigma+\kappa U_0)\left(\frac{p_1}{\rho}-2\nu\frac{dw_1}{dz}\right) = 0 \quad (z = 0), \tag{10·32}$$

$$g\zeta_1 = \frac{p_1}{\rho}-2\nu\frac{dw_1}{dz} \quad (z = 0). \tag{10·33}$$

In order to exclude surface waves to the first order, we must have

$$\zeta_1 = 0, \tag{10·41}$$

and then from (10·33) we see that

$$\frac{p_1}{\rho} - 2\nu\frac{dw_1}{dz} = 0 \quad (z = 0). \tag{10·42}$$

The condition (10·42) satisfies (10·32), but we may also have $\sigma + \kappa U_0 = 0$.

From (10·23) and (10·31) we deduce

$$\frac{d^2 w_1}{dz^2} = 0 \quad (z = 0); \tag{10·51}$$

from (10·24), (10·31) and (10·42) we deduce

$$\left\{\nu\left(\frac{d^2}{dz^2} - 3\kappa^2 - 3k^2\right) - i(\sigma + \kappa U_0)\right\}\frac{dw_1}{dz} = 0 \quad (z = 0), \tag{10·52}$$

and from (10·23) and (10·31)

$$\frac{d\chi_1}{dz} = 0 \quad (z = 0). \tag{10·53}$$

The solution of the equation (10·21) involves four constants of integration in w_1, three of these are required to satisfy (10·31), (10·51) and (10·52), while the fourth determines the order of magnitude of w_1. The solution of the equation (10·22) involves two constants of integration in χ_1; one of these is required to satisfy (10·53) and the other determines the magnitude of χ_1. We therefore have two constants of integration, one in w_1 and the other in χ_1, with which to satisfy the two conditions (7·42) and (7·43) at $z = -h$.

When the disturbance is two-dimensional and confined to the plane of x and z, so that $k = 0$ and $v_1 = 0$, we have also $\chi_1 = 0$. The available constants of integration are then reduced to one and it is impossible for one solution of the type (10·1) to satisfy both the conditions (7·42) and (7·43).

On using accents to denote conjugate complexes and utilizing such relationships as

$$[\cos^2(\sigma t + \kappa x)\sin^2 ky] = \tfrac{1}{4},$$

we have

$$[uw] = n(u_1 w_1' + u_1' w_1),$$

$$[\tfrac{1}{2}q^2 + w^2] = n(u_1 u_1' + v_1 v_1' + 3w_1 w_1'),$$

$$[w\,.\,\tfrac{1}{2}q^2] = 0,$$

$$[wp] = n(w_1 p_1' + w_1' p_1),$$

where $n = \frac{1}{3}$ when $k \neq 0$, and $n = \frac{1}{4}$ when $k = 0$. From (10·23) we deduce

$$(\kappa^2 + k^2)(u_1 u_1' + v_1 v_1') = \frac{dw_1}{dz}\frac{dw_1'}{dz} + \chi_1 \chi_1'.$$

When we substitute into the right-hand sides of the equations of § 7 the real parts of (10·1), and then reintroduce complex quantities, we find that t, x and y enter (7·21) through the factors

$$e^{2i(\sigma t + \kappa x)} \cos 2ky, \quad e^{2i(\sigma t + \kappa x)}, \quad \cos 2ky,$$

and (7·22) through the factors

$$e^{2i(\sigma t + \kappa x)} \sin 2ky, \quad \sin 2ky.$$

The second approximation will therefore contain three new groups of terms, each of the same general form as (10·1).

11. UNIFORM MEAN CURRENT

We now take $U = U_0$, when U_0 is a constant, and write

$$m^2 = \kappa^2 + k^2.$$

We also take a and b to denote constant speeds which fix the magnitude of the disturbance.

We first examine the case in which $\sigma + \kappa U_0 = 0$, so that the disturbance is convected by the mean current. Then a solution of the equations of § 10 is given by

$$u = -\sin \kappa(x - U_0 t) \cos ky \left\{ \frac{\kappa a}{m} (mz \sinh mz + \cosh mz) - \frac{kb}{m} \cosh mz \right\},$$

$$v = -\cos \kappa(x - U_0 t) \sin ky \left\{ \frac{ka}{m} (mz \sinh mz + \cosh mz) + \frac{\kappa b}{m} \cosh mz \right\},$$

$$w = amz \cos \kappa(x - U_0 t) \cos ky \cosh mz,$$

$$\frac{p}{\rho} = 2\nu ma \cos \kappa(x - U_0 t) \cos ky \cosh mz,$$

and of course this does not reduce to the first example of § 9 when we write $\nu = 0$. $k = 0$.

The second-order terms take the form

$$u = \frac{1}{\nu}\sin 2\kappa(x-U_0t)\{u_2\cos 2ky + u_2'\},$$

$$v = \frac{1}{\nu}\{v_2\cos 2\kappa(x-U_0t)+v_2''\}\sin 2ky,$$

$$w = \frac{1}{\nu}\{w_2\cos 2\kappa(x-U_0t)\cos 2ky + w_2'\cos 2\kappa(x-U_0t)+w_2''\cos 2ky\},$$

$$\frac{p}{\rho} = p_2\cos 2\kappa(x-U_0t)\cos 2ky + p_2'\cos 2\kappa(x-U_0t)+p_2''\cos 2ky,$$

where $u_2, u_2', \ldots, p_2''$ are functions of z, quadratic in a and b, and independent of ν.

We deduce that

$$[vw] = [wu] = [uv] = 0, \quad [w\,.\,\tfrac{1}{2}q^2] = 0,$$

$$\begin{aligned}[\tfrac{1}{2}q^2+w^2] = {} & \tfrac{1}{8}a^2\{m^2z^2(2\cosh 2mz+1)+mz\sinh 2mz\} \\ & +\tfrac{1}{16}(a^2+b^2)(\cosh 2mz+1) \\ & +\frac{1}{8\nu^2}\{u_2^2+v_2^2+2(u_2'^2+v_2''^2)+6w_2^2+12(w_2'^2+w_2''^2)\},\end{aligned}$$

$$\frac{1}{\rho}[wp] = \tfrac{1}{4}\nu ma^2\,.\,mz(\cosh 2mz+1)+\frac{1}{4\nu}(w_2\,p_2+2w_2'\,p_2'+2w_2''\,p_2'').$$

We then see that the condition (7·42) cannot be satisfied with a non-zero value of T, and that the energy equation (5·2), with $z_2 = 0$, reduces to

$$\nu\int_{z_1}^{0}[\phi]\,dz = \frac{1}{\rho}[wp]-\nu\frac{d}{dz}[\tfrac{1}{2}q^2+w^2],$$

where the right-hand side is evaluated at $z = z_1$.

Though this example does not provide forces of internal friction, it does illustrate the doing of work by turbulent pressures. We conclude, however, that where internal friction occurs, the disturbance cannot be completely convected by the mean current.

We next examine the case in which $\sigma+\kappa U_0 \neq 0$, and write

$$R = \frac{\sigma+\kappa U_0}{2m^2\nu}, \quad \lambda^2 = 1+2iR.$$

Then the equations of § 10 are satisfied by

$$u_1 = \frac{i\kappa a}{m}\{\cosh mz - \tfrac{1}{2}(\lambda^2+1)\cosh\lambda mz\} + \frac{kb}{m}\cosh\lambda mz,$$

$$v_1 = -\frac{ka}{m}\{\cosh mz - \tfrac{1}{2}(\lambda^2+1)\cosh\lambda mz\} - \frac{i\kappa b}{m}\cosh\lambda mz,$$

$$w_1 = a\left\{\sinh mz - \frac{1}{2}\,\frac{\lambda^2+1}{\lambda}\sinh\lambda mz\right\},$$

$$\frac{p_1}{\rho} = m\nu a(1-\lambda^2)\cosh mz.$$

The second-order terms have the general character indicated in § 10. The coefficient of viscosity ν does not now enter in the same simple way as when $\sigma+\kappa U_0 = 0$. On account of the complication of these terms we shall not give them.

We shall now write $\qquad 2\lambda = \alpha + i\beta,$

where α and β are real, so that

$$\tfrac{1}{2}\alpha^2 = (1+4R^2)^{\frac{1}{2}}+1, \quad \tfrac{1}{2}\beta^2 = (1+4R^2)^{\frac{1}{2}}-1.$$

In the above a and b may be complex, but in what follows we shall, for simplicity, suppose them to be real. On utilizing such relationships as

$$[\cos\tfrac{1}{2}\beta mz] = 0, \quad [\cos^2\tfrac{1}{2}\beta mz] = \tfrac{1}{2},$$

in addition to those used near the end of § 10, we have

$$[uv] = [vw] = 0, \quad [w\,.\,\tfrac{1}{2}q^2] = 0, \quad [wp] = 0,$$

$$[uw] = -\frac{\kappa a\beta(1+R^2)+kb(\alpha+R\beta)}{16m(1+4R^2)^{\frac{1}{2}}}\,a\sinh\alpha mz,$$

$$[\tfrac{1}{2}q^2+w^2] = \tfrac{1}{8}a^2(2\cosh\alpha mz - 1)$$
$$+\tfrac{1}{16}\{(1+R^2)(1+3(1+4R^2)^{-\frac{1}{2}})a^2+b^2\}\cosh\alpha mz.$$

We can now choose a and b so as to satisfy (7·42) and (7·43) at $z = -h$.

At last we have got a particular example in which $[uw]$ is not zero, and in which all the conditions we have prescribed are satisfied.

We notice, from (3·5), that the eddy-viscosity function N is infinite.

The presence of the hyperbolic functions in the above formulae is, of course, due to the uniformity of U. To obtain a case more like that of actual tidal currents we should require U to be such a function of z that $[uw]$ was proportional to z. The investigation of such a function would require more elaborate mathematical methods than those used in this paper.

REFERENCES

Godske, C. L. 1934 *Astrophys. Norvegica*, **1**, 11–86.
Karman, T. de and Howarth, L 1937 *Proc. Roy. Soc.* A, **164**, 192–215.
Reynolds, Osborne 1894 *Phil Trans.* A, **186**, 123–164.
Taylor, G. I. 1935 *Proc. Roy. Soc.* A, **151**, 421–478.

The kinetics of the thermal decomposition of acetophenone

BY R. E. SMITH AND C. N. HINSHELWOOD, F.R.S.

(*Received* 15 *July* 1940)

The mechanisms involved in the thermal decomposition of aldehydes and ketones are varied, and the relations between them somewhat complex. In particular, an interesting contrast in behaviour has recently been found between benzaldehyde and acetaldehyde. An investigation of acetophenone has therefore been made for comparison with acetone.

The thermal decomposition of acetophenone takes place predominantly by the step $C_6H_5COCH_3 = C_6H_5CH_3 + CO$, the toluene undergoing a subsequent decomposition to give chiefly benzene, methane and carbon. It differs from that of acetone in yielding hardly any ketene. It is homogeneous and nearly of the first order, with no sharp falling off in rate at 20 mm. There is no retardation by nitric oxide or by greatly increased surface, nor can an increased rate of decomposition be induced by the presence of radicals from decomposing diethyl ether. This is taken as evidence for the absence of reaction chains. In this respect the behaviour resembles that of acetone, but other differences in kinetics exist. These are briefly discussed.

INTRODUCTION

The homogeneous thermal decompositions of aldehydes and ketones appear to be of two types. One, generally accepted to depend upon a chain reaction, is inhibited by nitric oxide and occurs with benzaldehyde and with the higher aliphatic aldehydes: the other is not inhibited by nitric oxide and is believed by some, for this and other reasons, to depend upon an internal molecular rearrangement. It occurs as the sole type of mechanism with formaldehyde, acetaldehyde, chloral, acetone and methyl ketone, and accompanies the chain reaction with the compounds mentioned above.

The products are in general carbon monoxide and a molecule formed from the two residual radicals, though with acetone large quantities of ketene are formed and undergo secondary decomposition.

The reaction of formaldehyde is of the second order, that of acetone and that of benzaldehyde nearly of the first: acetaldehyde and propaldehyde seem to show a superposition of two or more mechanisms.

The variety and complexity of behaviour, the unsettled differences of opinion about mechanism, and the unexpected contrast recently found between benzaldehyde and acetaldehyde (Smith and Hinshelwood 1940, other references here) provide reasons for filling in an existing gap, namely, the comparison of acetone and acetophenone, by investigating the kinetics of the reaction of the latter.

The experimental method was the same as that used in the recent study of benzaldehyde (Smith and Hinshelwood 1940).

COURSE OF THE REACTION

By analogy with acetone, the intermediate production of ketene might be expected: $C_6H_5COCH_3 = C_6H_6 + CH_2 \cdot CO$. The amount of ketene was measured by withdrawing samples in a gas pipette, shaking with water and titrating with sodium hydroxide.

Since ketene decomposes at approximately the same rate as acetophenone at equal concentration, its formation would lead to the building up of a considerable stationary concentration The amount formed is, however, at all times very small (table 1), and hence the proportion of acetophenone giving ketene as an intermediate product must also be small. This conclusion is verified by the analyses of the permanent gases (table 2), in which the percentage of carbon monoxide goes down and not up as the reaction proceeds.

TABLE 1

Pressure increase in mm.	25	60	90	110	150
Pressure of ketene in mm	2·1	3·1	4·0	1·6	1·1

Since the proportions of the products change during the course of the reaction there must be some secondary decomposition. From table 2 it can be seen that the ratio CH_4/CO tends to a very small value in the initial stages, so that the methane must be formed by the decomposition of the intermediate product. Since this product is not in the permanent gas it must be a benzene compound and is presumably toluene. Separate experiments have shown that toluene does decompose at this temperature, 605° C (at a rate approximately twice that of acetophenone at equal concentration) to give the same products, methane, hydrogen and ethylene in the proportions 70 : 19 : 11, which are nearly the same as those found in the secondary decomposition with acetophenone.

TABLE 2

% decomposed	% CO	% CH_4	Ratio CH_4/CO	% C_2H_4	% H_2	% CO_2
19	78·6	13·2	0·168	3·7	6·5	1·8
31	72·7	18·3	0·252	3·9	3·5	1·6
53	71·9	18·9	0·263	2·8	4·2	2·2
71	67·1	25·4	0·378	2·1	4·1	1·3
92	63·6	27·9	0·439	1·8	6·0	0·7

Since the ratio CH_4/CO extrapolated to the beginning of the reaction is less than one-fifth of the final value, at least four-fifths of the acetophenone must yield toluene as an intermediate product. Hence the predominant course of the reaction seems to be.

$$CH_3COC_6H_5 = CO + CH_3C_6H_5$$

$$2CH_3C_6H_5 = 2C_6H_6 + CH_4 + C$$

This hypothesis is further substantiated by the following observations:

(1) The ratio CH_4/CO tends to 1/2 at the end of the reaction.

(2) Carbon is deposited on the walls of the reaction vessel.

(3) The pressure of the final products is 235–240 % of the initial pressure of acetophenone. The defect from the theoretical value of 250 % is to be expected from the dead space of the system.

The fraction of high boiling products can be determined by measuring the deviation from Charles's law on cooling. Approximately 2·5 % of diphenyl, at the end point, estimated by this method is probably due to a subsequent reaction of the benzene.

The possibility of changes in mechanism in the experimental ranges of temperature and pressure could not be neglected. Hence pressure-time curves for initial pressures of acetophenone from 50 to 150 mm. and for temperatures from 570 to 630° C were compared on adjusted time scales. They were found to be practically superimposable throughout their course so that no change in the nature of the reaction is indicated.

As a secondary decomposition exists, the pressure increase does not give, throughout the course of the reaction, an exact indication of the amount of acetophenone decomposing. For this reason the initial rate of pressure increase has been used as the standard of rate. Actually, it was found that comparisons made by this standard did not differ much from those based upon the reciprocal of the time for a pressure increase of 25 or 50 % of the initial pressure.

KINETICS

The thermal decomposition of acetophenone is homogeneous at 605° C and is very nearly of the first order (table 3).

TABLE 3

Unpacked reaction vessel		Packed reaction vessel	
P_0	$1/p(dp/dt) \times 10^3$	P_0	$1/p(dp/dt) \times 10^3$
19	1·62	26	1·61
40	1·83	64	1·85
63	1·84	63	1·79
87	1·95	94	1·94
120	2·4	101	2·30
131	2·1	111	2·14
172	2·28	136	1·99
182	2·3	139	2·08
234	2·35	177	2·33
274	2·49	203	2·24
283	2·7	243	2·46

The activation energy, determined for an initial pressure of 150 mm. (table 5), is 70,000 cal. per g.mol. (by least squares).

TABLE 4

Pressure NO mm.	$1/p(dp/dt) \times 10^3$
0·0	1·87
0·2	1·88
0·4	1·91
1·0	1·98
2·0	1·83
5·0	1·89
10	1·87
17	1·81
24	2·08
30	2·05

No evidence for a chain reaction was found. The reaction is neither inhibited nor appreciably accelerated by the addition of nitric oxide (table 4). No retardation (other than 4 % expected from the relatively increased dead space) was found with a reaction vessel packed with small silica spheres, the much larger surface of which would be expected to break a fraction of any chains, especially at low pressures. Since the reaction is of the first order and not of the 3/2 order, a hypothetical chain mechanism must, as with benzaldehyde, involve the unimolecular decomposition of one of the participating radicals. For example:

$$\begin{aligned} CH_3COC_6H_5 &= C_6H_5 + CH_3CO \\ CH_3CO &= CH_3 + CO \\ CH_3 + CH_3COC_6H_5 &= CH_3C_6H_5 + CO + CH_3 \\ CH_3 + CH_3CO &= C_2H_6 + CO \end{aligned}$$

TABLE 5

Temperature °C	$1/p(dp/dt) \times 10^3$
635·5	9·8
616	4·77
605	2·19
590	1·25
575	0·66

TABLE 6

Time min.	(1) Δp 40 mm. ether	(2) Δp 100 mm. acet.	(1)+(2)	Δp mixture
2	13	2	15	19
4	25	4	29	34
6	36	6	42	47
8	45	8	53	56
10	51	11	62	64
12	56	13	69	71
14	61	15	76	77
16	64	17	81	82
18	67	19	86	86
20	69	20	89	89

The decomposition of benzaldehyde, depending upon rather similar processes, is inhibited by nitric oxide. Such a mechanism, if it existed with acetophenone, would also be expected to suffer inhibition.

The decomposition might involve free radicals, but in a non-cyclic set of reactions, though this does not seem likely since no ethane and little diphenyl were detected in the products. Decomposing diethyl ether, a known source of radicals, was mixed with acetophenone at 555° C; where the latter decomposes very slowly. Table 6 shows that the interaction does not increase the rate of pressure change. Hence, we conclude that methyl radicals, whether they occur free in the normal decomposition or not, could hardly stimulate a chain reaction in the acetophenone. Thus it seems probable that the absence of chains in acetophenone could be explained not only by the lack of initiating radicals but also by the non-existence of a suitable set of recurring reactions.

The contrast in behaviour between acetaldehyde and benzaldehyde is not found with the corresponding pair acetone and acetophenone, neither of which show any evidence of giving a chain reaction. Two differences do, however, appear: acetone gives much more ketene than acetophenone; and the rate-initial pressure curves for the two substances, although both approximating to those of first order reactions, differ in detail. With acetophenone the curve does not become quite parallel to the pressure axis at the higher pressures, as with acetone, nor does it show the same rather rapid fall below 100 mm. These differences would be consistent with the theory of multiple modes of activation (Hinshelwood, Fletcher, Verhoek and Winkler 1934), though other explanations may exist.

At 605° C and 100 mm. the value of $1/p(dp/dt)$ for acetophenone is $2{\cdot}0 \times 10^{-3}$: and for acetone $8{\cdot}5 \times 10^{-3}$, the difference being accounted for within the limits of accuracy by the change in activation energy. Those aldehyde decompositions which, according to our present view, are chain-free have considerably smaller activation energies than the ketone reactions; the absolute reaction rates are not correspondingly greater. The non-exponential term of the Arrhenius equation diminishes steadily from the ketones through benzaldehyde and the higher aliphatic aldehydes to formaldehyde, the trend being explained, according to the view referred to above, by the progressive simplification through this series of the internal activation mechanism governing the reaction.

REFERENCES

Hinshelwood, Fletcher, Verhoek and Winkler 1934 *Proc. Roy. Soc.* A, **146**, 327.
Smith and Hinshelwood 1940 *Proc. Roy. Soc.* A, **175**, 131 (other references here).

The poisoning of a palladium catalyst by carbon monoxide

By M. G. T. Burrows and W. H. Stockmayer
Leoline Jenkins Laboratories, Jesus College, Oxford

(*Communicated by D. L. Chapman, F.R.S.—Received* 29 *April* 1940)

The combination of hydrogen and oxygen on the surface of a palladium catalyst at low partial pressures at the temperature of the laboratory is completely inhibited by very small amounts of carbon monoxide. The poisoning effect is only temporary and, during the induction periods, which are observed, the carbon monoxide is removed by the oxygen in the mixture of hydrogen and oxygen. After the induction period combination proceeds at the normal rate. The length of the inert periods observed increases with the amount of carbon monoxide initially added to the system, and the rate of removal of the carbon monoxide increases as its pressure decreases, and becomes comparatively very great when the carbon monoxide present does not exceed the amount required to cover the surface of the metal.

The observations recorded can be explained by a development of the hypothesis of De la Rive that the catalysis of the reaction between hydrogen and oxygen, and between carbon monoxide and oxygen, involves the alternate oxidation of the metal and the reduction of the surface oxide by hydrogen or by carbon monoxide respectively.

In 1934 Chapman and Gregory investigated the kinetics of the interaction of hydrogen and oxygen at low partial pressures of the gases and at the temperature of the laboratory. The results supported the hypothesis of De la Rive and Becquerel (1839), namely, that the production of water vapour was due to the alternate oxidation of the palladium and reduction of the oxide, formed on the surface, by hydrogen.

As we shall show the kinetics of the interaction of carbon monoxide and oxygen also support the mechanism suggested by De la Rive, and recently one of us (M G. T. B.) has obtained independent and almost unequivocal evidence (which will be subsequently published) of the truth of the hypothesis.

Chapman and Gregory observed incidentally, during the course of their research, a striking phenomenon which was explained by a working hypothesis shown subsequently by Cadwallader (1935) to be untenable. Nevertheless this hypothesis, although inapplicable to the case of hydrogen and oxygen, is almost certainly applicable to that of carbon monoxide and oxygen.

It is important that the observations of Cadwallader should be briefly described, since they demonstrate the supreme importance of removing from the internal surface of the glass apparatus traces of adsorbed and dissolved gas in work of this class. He describes a simple method of effecting this removal in cases where, owing to the presence of mercury and glass taps, the usual process of complete 'outgassing' is impracticable.

Chapman and Gregory observed that, when the palladium was heated for a long time in hydrogen, it was rendered inert as a catalyst for the reaction between hydrogen and oxygen, but that on being allowed to stand for a long time in the presence of these gases, it suddenly and completely recovered its catalytic activity. Furthermore, the palladium which had been heated in hydrogen and then in a good vacuum behaved in the same manner. It was observed that the palladium could be reactivated by its exposure to oxygen.

Cadwallader, with an apparatus which had been constructed and cleaned with the usual precautions, repeated and confirmed the above observations of Chapman and Gregory. But, as his results were irregular, he decided to try the effect of passing the silent discharge alternately through hydrogen and oxygen contained in the apparatus. During the passage of the discharge through both gases condensible vapours collected in the liquid air traps, at first rather rapidly, but after several days' treatment, very slowly. After this treatment of the apparatus induction periods were no longer observed after the catalyst had been heated in hydrogen.

He therefore surmised that carbon dioxide adsorbed by the glass gradually escaped; was not entirely held back in the liquid air trap—its vapour pressure at the temperature of liquid air is approaching 10^{-6} mm. of mercury—and was reduced to the poison carbon monoxide by hydrogen in contact with the red-hot palladium. This was confirmed by showing that carbon monoxide, at a partial pressure too small to be measured with the McLeod gauge, in electrolytic gas caused induction periods of the same character as those observed by Chapman and Gregory. It was, of course, necessary to assume that carbon monoxide is removed from the surface of the palladium and from the apparatus during the induction period by the oxygen in the electrolytic gas.

In the present communication an account is given of an investigation of the poisoning effect of carbon monoxide on the hydrogen-oxygen reaction at a palladium surface.

EXPERIMENTAL

The apparatus was similar to that employed by Chapman and Gregory. It consisted essentially of a reaction vessel protected by two liquid air traps; six gas reservoirs; a pressure adjuster and sources of hydrogen, oxygen and carbon monoxide. Pressure measurements were made on a McLeod gauge. Exhaustion was effected by a pumping system consisting of a mercury vapour pump backed by an automatic Sprengel pump. The various units were connected to the leads through mercury seals.

The reaction vessel, which was made of quartz, was connected to the rest of the apparatus by means of a ground joint lubricated with a mixture of meta- and orthophosphoric acids. A glass tap, lubricated with the same mixture, as described by Chapman and Moignard (1937), was inserted between the two liquid air traps.

The catalyst, which lay in the reaction vessel, was a thin strip of palladium foil 11 mm. long, 1 mm. wide and 0·03 mm. thick.

Two capillary reservoirs of known volume were used for the isolation of small amounts of carbon monoxide. The gas enclosed in them could be let into the leads, and from thence into the reaction vessel by sucking down the mercury which closed their entrances.

Hydrogen was prepared in a pure state by diffusion of hydrogen, prepared in a Kipp's generator from arsenic-free zinc and pure sulphuric acid, through a heated palladium thimble.

Oxygen was prepared by heating dry recrystallized potassium permanganate. The oxygen was collected in a large reservoir filled with sticks of potash, and it was admitted to the apparatus through a U-tube surrounded with liquid air

A pressure adjuster was used to mix the gases in the correct proportions. The mixtures were stored in the gas reservoirs.

Carbon monoxide was prepared by heating an intimate mixture of sodium formate and metaphosphoric acid. Purification was effected by leaving over potash and cooling to the temperature of liquid air.

After evacuation of the apparatus the interior of the glass was cleaned by passing an electric discharge alternately through oxygen at 0·3 mm. pressure and hydrogen at 1 mm. pressure contained in the apparatus. This process eliminated impurities which, if not removed in the trap surrounded with liquid air, would have poisoned the catalyst. The palladium was rendered active by its being heated in oxygen at 700° C.

After the carbon monoxide had been shut up in the capillary reservoirs at a known pressure, the leads were evacuated and thoroughly washed out with electrolytic gas. On letting the carbon monoxide out of either of these reservoirs, its pressure was too small to be measured on the McLeod gauge; but it could be calculated, since the ratio of the volume of the capillary reservoir to the leads was known.

Sufficient carbon monoxide to give a monomolecular layer of adsorbed carbon monoxide on the palladium catalyst would give a pressure of $8{\cdot}3 \times 10^{-6}$ mm. in the leads, or $5{\cdot}2 \times 10^{-6}$ mm. in the leads and reaction vessel. For the purposes of calculation it was assumed that there was a chequer-board arrangement of tangent carbon monoxide molecules on the surface of the palladium, so that the area covered per molecule would be the square of the molecular diameter. From published data Jeans (1916) gives $3{\cdot}78 \times 10^{-8}$ cm. as the best value for the diameter of a carbon monoxide molecule. From this value it was calculated that $1{\cdot}6 \times 10^{14}$ molecules would completely cover the surface of the catalyst which was 22·7 sq. mm. in area.

The induction periods observed after exposure of the palladium catalyst to carbon monoxide

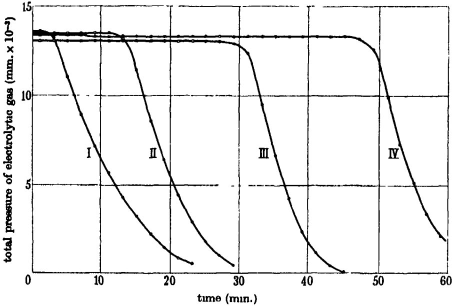
In the majority of the experiments the catalyst was exposed to carbon monoxide for a fixed time, electrolytic gas was then introduced and pressure readings taken at frequent intervals.

In a typical experiment the catalyst was exposed to carbon monoxide for 3 hr. Electrolytic gas was then added at a pressure of 0·1 mm. and pressure readings taken every 2 min. No fall in pressure took place for 15 min., then reaction began and proceeded at the normal rate. The partial pressure of the carbon monoxide in the leads and reaction vessel was $4{\cdot}4 \times 10^{-5}$ mm. corresponding to 8·5 'layers'. The results of other experiments are plotted in the diagram below.

That the carbon monoxide was removed from the system by the oxygen present in the electrolytic gas was demonstrated by the following experiments. Mixtures of a small amount of carbon monoxide and oxygen, and of a small amount of carbon monoxide and hydrogen, were shut up in the reaction vessel and hydrogen and oxygen respectively admitted to the leads. After leaving for a suitable time the seal to the reaction vessel was lowered and pressure measurements taken in the normal manner.

Induction periods were only observed when the mixture enclosed in the reaction vessel was carbon monoxide and hydrogen. When the mixture enclosed in the reaction vessel was carbon monoxide and oxygen all the carbon monoxide had combined with the oxygen before the hydrogen was added.

When the constituents of electrolytic gas combine in the presence of the catalyst its surface is liable to be oxidized after the reaction has stopped. On exposure of the metal to carbon monoxide some of the carbon monoxide is removed in reducing this oxide and, in consequence, the length of the induction period is lower than it would have been if the surface of the metal had been unoxidized. Therefore the palladium was always exposed to cold hydrogen before carrying out an experiment to measure the length of an induction period. That superficial oxidation of the catalyst occurs when it is exposed to oxygen for a short time is demonstrated by the experiments quoted below. The pressure of the electrolytic gas used was 0·1 mm.



Graphs of induction periods with varying initial amounts of carbon monoxide.

Experiment I. Initial pressure of carbon monoxide = 0·78 × 10^{-5} mm., corresponding to 1·5 layers. Induction period = 2·5 min.
Experiment II. Initial pressure of carbon monoxide = 1·56 × 10^{-5} mm., corresponding to 3 layers. Induction period = 13 min.
Experiment III. Initial pressure of carbon monoxide = 2·6 × 10^{-5} mm., corresponding to 5 layers. Induction period = 30·5 min.
Experiment IV. Initial pressure of carbon monoxide = 3·64 × 10^{-5} mm., corresponding to 7 layers. Induction period = 48·5 min.

Catalyst exposed to oxygen for 1 hr., oxygen pumped off, then catalyst exposed to carbon monoxide (pressure 3·3 × 10^{-5} mm.).	Catalyst exposed to hydrogen for 1 hr., hydrogen pumped off, then catalyst exposed to carbon monoxide (pressure 3·2 × 10^{-5} mm.).
Induction period = 2·9 min.	Induction period = 10·3 min.

The length of the inert period induced by exposure of the palladium to carbon monoxide increases with the amount of carbon monoxide introduced. The rate of removal of carbon monoxide decreases as its pressure increases. When the amount of carbon monoxide is small the rate of its removal becomes comparatively very great. This is indicated by the sharp ending of the induction periods.

At first some difficulty was experienced in obtaining reproducible induction periods due to the adsorption of carbon monoxide on the walls of the traps cooled to the temperature of liquid oxygen. We found, however, that this source of error could be eliminated by covering the cooled surface with a thin layer of ice, on which carbon monoxide is not appreciably adsorbed. The covering of the cooled surface with ice was easily effected by the condensation of water vapour formed in the reaction vessel from oxygen and hydrogen.

The smallest amount of carbon monoxide which can entirely prevent reaction was found to be of the order of a monomolecular layer. This indicates that, to deprive the metal completely of its catalytic activity, the whole of the palladium surface must be covered with carbon monoxide.

In the experiments quoted below the catalyst was exposed to the carbon monoxide for ½ hr. and then electrolytic gas let into the reaction vessel at a pressure of 0·02 mm. The temperature during the experiments was 20° C.

Number of experiment	Pressure of CO used in initial exposure mm.	Induction period in min.		
		Observed	Calculated using equation (1) given below	Calculated using equation (2) given below
1	$1{\cdot}924 \times 10^{-5}$	3·5	3·5	3·2
2	$2{\cdot}444 \times 10^{-5}$	5	5·1	4·8
3	$2{\cdot}912 \times 10^{-5}$	6·5	6·8	6·4
4	$3{\cdot}693 \times 10^{-5}$	10·5	10·0	9·7
5	$4{\cdot}68 \times 10^{-5}$	13·75	15·1	14·8
6	$5{\cdot}406 \times 10^{-5}$	17·75	19·4	19·3
7	$5{\cdot}954 \times 10^{-5}$	24·25	23·2	23·2
8	$6{\cdot}76 \times 10^{-5}$	26	29·0	29·1
9	$8{\cdot}84 \times 10^{-5}$	52	47·8	48·1
10	$13{\cdot}52 \times 10^{-5}$	107·5	108·8	108·2
11	$13{\cdot}52 \times 10^{-5}$	113	108·8	108·2
12	$22{\cdot}05 \times 10^{-5}$	294·5	294·5	280·0

The times of the induction periods were calculated from the power series

$$t = a + bp + cp^2 + dp^3, \tag{1}$$

where t is the time in minutes and p the pressure. The constants were calculated by the method of least squares and found to be

$$a = -0{\cdot}07, \quad b = 9{\cdot}44 \times 10^4, \quad c = 4{\cdot}655 \times 10^9, \quad d = 4{\cdot}42 \times 10^{12}.$$

Attention must be drawn to the fact that the above empirical formula cannot be used for the purpose of extrapolation to zero time, since the constants a and b are appreciably affected by the results of the experiments performed with moderate and high initial pressures of carbon monoxide.

For initial pressures of carbon monoxide exceeding $2{\cdot}5 \times 10^{-5}$ mm. the simpler empirical formula

$$t = ap + bp^2, \tag{2}$$

in which $a = 6{\cdot}04 \times 10^4$ and $b = 5{\cdot}474 \times 10^9$ was found to be in good agreement with the experimental results.

The decrease in the rate of removal of carbon monoxide as its pressure increases was also shown in numerous other series of determinations.

A summary of another typical set of observations carried out at 20° C using a lower pressure (0·006 mm.) of electrolytic gas during the induction period is given below.

Number of experiment	Initial pressure of carbon monoxide in the leads and reaction vessel mm.	Pressure difference mm.	Induction period in min.	Difference in the induction periods
1	$0{\cdot}52 \times 10^{-5}$		2·5	
2	$1{\cdot}04 \times 10^{-5}$	$1{\cdot}04 \times 10^{-5}$	6	7
3	$1{\cdot}56 \times 10^{-5}$	$1{\cdot}04 \times 10^{-5}$	9·5	11·5
4	$2{\cdot}08 \times 10^{-5}$	$1{\cdot}04 \times 10^{-5}$	17·5	20·5
5	$3{\cdot}12 \times 10^{-5}$	$1{\cdot}04 \times 10^{-5}$	38	28
6	$4{\cdot}16 \times 10^{-5}$	$1{\cdot}04 \times 10^{-5}$	66	27
7	$5{\cdot}2 \times 10^{-5}$		93	

It will be seen that at the lowest pressure of carbon monoxide an increase of pressure of $1{\cdot}04 \times 10^{-5}$ mm. caused an increase in the induction period of 7 min., whereas at the highest pressure of carbon monoxide an increase of pressure of the same magnitude resulted in an induction period 27 min. longer.

To confirm the above conclusion a number of experiments were carried out in which the catalyst was exposed to a constant amount of carbon monoxide, but, in certain cases, the partial pressure of carbon monoxide was reduced. This was effected by opening one or more of the gas reservoirs.

The decrease in the partial pressure of carbon monoxide due to the increase in volume reduced the length of the resulting induction period.

The pressure of electrolytic gas during the induction period was kept constant at 0·02 mm. and the temperature maintained at 20° C.

Number of experiment	Volume	Induction period in min.
1	1	60·5
2	1	56·6
3	2	50·2
4	1	66·2
5	1	68·5
6	1	61·8
7	1	61
8	2	53
9	1	64
10	2	55·5
11	1	65·7
12	2	62
13	1	182
14	2	139
15	1	183·4
16	2	154·2

In the experiments quoted above the initial amount of carbon monoxide was kept constant during the course of each group of determinations.

Thus it was found that an increase in the pressure of the electrolytic gas resulted in a decrease in the length of the induction period. Furthermore, with mixtures of the gases not in combining proportions, if the partial pressure of the oxygen was kept constant, the length of the inert period induced by exposure to carbon monoxide was independent of the partial pressure of hydrogen If, however, the partial pressure of oxygen was increased, that of the hydrogen being kept constant, the length of the induction period was considerably diminished. These results prove that it is the oxygen in the electrolytic mixture which removes the carbon monoxide from the system. (See table below.)

Variations in temperature have a considerable effect on the length of the induction periods. Increasing the temperature reduces the length of the induction period as might be expected.

When mixtures of carbon monoxide and electrolytic gas were admitted to the reaction vessel (to which no carbon monoxide had previously been added) a small but measurable amount of combination occurred during the first minute but thereafter there was a definite induction period at the end of which reaction proceeded at the normal rate. This shows that a complete layer of carbon monoxide is formed on the catalyst surface in a very short time. Details of an experiment are given below: A mixture of $5{\cdot}564 \times 10^{-5}$ mm. of carbon monoxide and 0·1 mm. of electrolytic gas was let into the reaction vessel (temperature = 17·5° C). The pressure fall during the first minute was $5{\cdot}65 \times 10^{-4}$ mm.; thereafter reaction was prevented for 31 min. and then combination took place at a rate corresponding to a pressure fall of $25{\cdot}4 \times 10^{-4}$ mm. per min.

Number of experiment	Pressure of carbon monoxide initially present mm.	Partial pressure of hydrogen, partial pressure of oxygen being unity	Induction period in min.
1	$3\ 12 \times 10^{-5}$	2	30·5
2	$3\ 12 \times 10^{-5}$	1	28·5
3	$3{\cdot}12 \times 10^{-5}$	2	27·5
4	$3{\cdot}12 \times 10^{-5}$	1	32·5

Number of experiment	Pressure of carbon monoxide initially present mm.	Partial pressure of oxygen, partial pressure of hydrogen being unity	Induction period in min.
1	$3{\cdot}12 \times 10^{-5}$	¼	58
2	$3\ 12 \times 10^{-5}$	½	38·4
3	$3{\cdot}12 \times 10^{-5}$	¼	56
4	$3\ 12 \times 10^{-5}$	½	36 4

The results recorded in this communication can be explained by the hypothesis of De la Rive with one or two additional assumptions concerning the adsorptive properties of the metal.

As already stated, De la Rive postulated that the water formed when hydrogen and oxygen are brought into contact with the metal results from the reduction of the previously formed oxide. Similarly it must be assumed that the carbon dioxide produced when carbon monoxide and oxygen are brought into contact with the metal results from the reduction of the oxide of the metal by carbon monoxide.

To explain the observed phenomena we must also make the assumption that, when an amount of carbon monoxide a little greater than that required to form a monomolecular layer of the gas on the surface of the metal is introduced into the apparatus, it is nearly all adsorbed to form a layer of gas over the complete surface, so compact that only very occasionally is a molecule of oxygen able to come within the range of the valency forces of the metal so as to form with the latter palladium oxide. Accordingly both water vapour and carbon dioxide could be formed only very slowly. Moreover, as the pressure of the carbon monoxide in the apparatus is increased, the layer of carbon monoxide on the surface becomes more compact and the rate of formation of carbon dioxide is thereby diminished. Furthermore, when the carbon monoxide present is less than a monomolecular layer, its rate of removal by oxygen becomes very rapid owing to the fact that oxide can be formed on the surface left free, and consequently the rate of formation of water from electrolytic gas quickly attains its maximum value.

In conclusion we should like to thank Mr D L. Chapman, F.R.S., for his valuable advice and interest during the course of this investigation; and also to express our gratitude to the Department of Scientific and Industrial Research for a maintenance grant made to one of us (M.G.T.B.).

SUMMARY

1. Very small amounts of carbon monoxide completely inhibit the combination of hydrogen and oxygen at the surface of a palladium catalyst.

2. Carbon monoxide is a temporary poison. It is removed by the oxygen in the mixtures of hydrogen and oxygen.

3. When reaction begins, after an induction period, the maximum rate is rapidly attained.

4. The length of the induction period increases with increase in the amount of carbon monoxide admitted to the apparatus, but the rate of removal of carbon monoxide from the system decreases with increase in its pressure.

5. The length of the induction period decreases with increase in the pressure of electrolytic gas, and for non-electrolytic mixtures the length of the induction period is independent of the pressure of hydrogen, but decreases with increase in the pressure of oxygen.

6. The length of the inert period induced by a given amount of carbon monoxide decreases with rise in the temperature.

7. The observations recorded during the course of this investigation are in agreement with the theory that the catalysis is a process of alternate oxidation and reduction of the metal as first suggested by De la Rive.

REFERENCES

Cadwallader 1935 B.Sc thesis, Oxford
Chapman and Gregory 1934 *Proc. Roy. Soc.* A, **147**, 68.
Chapman and Moignard 1937 *J. Chem. Soc.* p. 1936.
De la Rive and Becquerel 1839 *Ann. Phys. u. Chem.* **46**, 489.
Jeans 1916 *The dynamical theory of gases*, ch. XIV, p. 341.

Investigations on the vibration spectra of certain condensed gases at the temperature of liquid nitrogen

I. Experimental methods

A. INFRA-RED SPECTRA

By G. K. T. Conn, E. Lee and G. B. B. M. Sutherland

Laboratory of Physical Chemistry, Cambridge

(*Communicated by R. G. W. Norrish, F.R.S.—Received* 10 *May* 1940)

A description is given of experimental methods which have been developed for the investigation of infra-red and of Raman spectra at the temperature of liquid nitrogen. The infra-red spectrometer is essentially the Pfund-Barnes type, but it has been found possible to dispense with the fore-prism and to separate the overlapping orders from the grating by means of filter shutters. The spectrometer is used as a monochromator so that the radiation only enters the absorption cell after it has been through the diffraction train. The absorption cell has been designed so that it may be used for high as well as low temperatures. The apparatus for the Raman effect is a modification of that introduced by De Hemptinne, being especially suitable for the examination of small quantities of material.

The infra-red spectrometer described here was designed with two main objects in view. The first was the investigation of molecules in the gaseous state, at ordinary temperatures, in order to determine their configuration, dimensions and internal force fields; the second was the study of thin layers of solidified gases at low temperatures, primarily in order to obtain evidence for or against the hypothesis of rotation in the solid state, but more broadly, to study the change in the spectrum which occurs with a change of state and at a transition point. The instrument is limited by the gratings at present available to the range from 1 to 25μ. It is based on the Pfund-Barnes (Pfund 1927; Barnes 1930) optical system, but there are several points in plan and technique of operation which we have developed for our own particular needs.

The plan of the instrument is illustrated in figure 1. The vertical Nernst glower, N, is focused on the entrance slit S_1 and the parabolic mirror M_2 produces a parallel beam which is dispersed by the grating G. Reflected rays falling on M_4 are brought to a focus at S_2 by a mirror M_5, similar to M_2. The plane slotted mirrors M_3 and M_4 are used so that radiation travels only along the axes of M_2 and M_5 thus minimizing distortion. By rotating the grating, the wave-length of the light focused on S_2 may be varied. The monochromatic radiation from S_2 is now made to converge and travel vertically downwards into the absorption cell by mirrors M_6 and M_7; from the cell it is reflected on to M_8, focused by M_9 on to the entrance slit of the thermocouple holder and finally on to the thermocouple by the short focus mirror M_{10}.

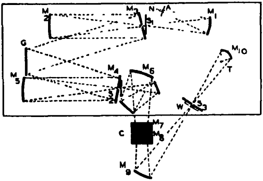


FIGURE 1. Optical arrangement of grating spectrometer.

The mirrors are all aluminized. The aperture ratio of M_2 and M_5 is 1/5 and their diameters 5 in. Up to the present, four gratings have been employed, having (*a*) 1200, (b_1) 4800, (b_2) 4800, (*c*) 7200 lines to the inch. The first two were ruled for us at the University of Michigan by Professor E. F. Barker, the former on tin and the latter on nickel. The other two are of a different type, for which we are indebted to Professor R. W. Wood of the Johns Hopkins University, Baltimore. These have been ruled on an aluminized glass flat and are much less expensive to make. With these gratings we have covered the spectrum from the visible to about 16μ with slit widths of the order of 1 cm.$^{-1}$, except when working far from the blaze angle of the grating, or in regions of strong atmospheric absorption.

In the first work carried out with the spectrometer (Conn and Sutherland 1939) a fore-prism was always used in front of the entrance slit S_1 to produce an initial low dispersion which separated fairly well the radiation under investigation from radiation of higher orders, i.e. shorter wave-length. For some of the later work (Lee and Wu 1939) a more convenient filter method has been employed. Though it is very difficult to find filters which are opaque to short waves and yet will transmit in longer wave-length regions, it is well known that certain easily obtainable materials will transmit all but a few per cent of the short waves and completely stop the longer waves. The incomplete transmission of the short-wave energy is due to the reflexions at the two surfaces of the filter and may be compensated for by introducing a second shutter (on removing the first) which reflects the same amount of the short-wave energy. Thus in the region of 3–4μ a Pyrex shutter (A in figure 1) 1 cm. thick used in conjunction with a glass cover-slip 0·15 mm. thick was found to be much more efficient than the fore-prism method. A gain in intensity of the order of 100% was achieved since losses due to absorption of the rock salt and to various additional reflexions had been eliminated. Difficulties due to poor dispersion of the rock-salt prism in the region are also obviated. For the region from 4–5μ we have used fused silica shutters of 6 and 2 mm. thickness together with a compensating plate of sylvine, and we hope to extend this method to cover the whole range of the instrument using various thicknesses of suitable materials.

Originally, the absorption cell followed the fore-prism in front of the entrance slit S_1, but the large amount of energy falling on the solidified gas caused heating and evaporation. In the present arrangement only the radiation of the particular wave-length where the absorption is being measured comes into the absorption cell and so such effects have been minimized. Again, when the fore-prism is used and a selected region of the prism spectrum is focused on S_1, the new arrangement has the additional advantage that many of the troublesome effects due to distortion of the slit image, produced in the absorption train, are avoided.

THE ABSORPTION CELL

The construction of an absorption cell for gases in the infra-red is simple and obvious and does not invite separate description so long as one is concerned with ordinary temperatures. As soon as we wish to raise or lower the temperature, appreciable difficulties arise. The type we used is shown in figure 2. It was primarily designed for the examination of solid layers at liquid air temperatures but it is equally suitable for gases at moderately high or low temperatures. It consists of a pyrex tube 7 cm. wide into one end of which a flat pyrex plate P has been sealed. This is silvered on the inside, the other end having a wide flange ground flat to take the cell window. The silvering is carried out by sputtering, using a discharge in argon at a suitable low pressure. The cell is inverted over a small brass table supporting insulated electrodes, the cathode consisting of a silver disc about half an inch from the pyrex flat. Apiezon Q compound is pressed round the flange and the cell is evacuated before letting in argon to carry the discharge. Using 3000 V between the electrodes a cathode dark space of about 1 cm. and a current of a few milliamperes, an opaque, highly reflecting mirror is obtained in about 10 min. Earlier attempts to get a good mirror by chemical methods were very unsatisfactory.

Owing to the size of the cell and the difficulty in obtaining large windows of rock salt, the following method was used to fix the window. A brass plate with a hole 4·5 cm. square cut in it is waxed on to the flange. The square rock-salt plate with a side of 5 cm. is then secured to the brass plate, porcenam paint being used to make the joint tight. The reasons for this composite window are (*a*) that it enables nearly the full aperture of the rock-salt window to be used and (*b*) that the window can be removed without injury to the rock-salt plate, as only the glass metal join is affected. There are two side limbs, the first to conduct the gas under examination and the second for the leads of a thermocouple which is hard soldered on to a piece of tungsten fused into the surface of the silvered base about 8 mm. from the side.* The advantage of this type of cell is that, being vertical, the lower part may readily be immersed in a constant temperature bath. Up to the present we have only used liquid air as the bath, or else worked with the cell at ordinary temperatures. In the former case, we have occasionally experienced trouble from the condensation of water vapour on the rock-salt window, but this has been overcome by having a small heating coil wrapped around the cell, just below the flange. In order to obtain 'cell out' readings for the estimation of percentage absorption, an additional plane mirror M_{11} is swung into the path of the beam. This mirror reflects the radiation on to an adjustable vertical mirror M_{12}, which returns it to the inclined mirror and into the original path of the beam from the cell. If desired, a compensating plate of the same material as that used for the cell window may be put in at R. The fact that this arrangement requires one window (or two at most) as against two (or four) of the conventional absorption train, is important, as such windows are expensive.

Figure 2 also shows the principle of the method used to control the level of the liquid air which is, of course, continually boiling away. The lower half of the Dewar flask contains a wide pyrex cylinder C open at the bottom and connected to the narrow vertical tube T. This can be raised or lowered by a simple ratchet arrangement which brings the orifice to any required distance below the surface of the oil in the cylinder G. A copper wire *W* runs from outside the Dewar flask into the liquid air in *C*. As the liquid air evaporates, a pressure is built up inside *C* which is controlled by the depth of the end of tube *T* below the oil surface. When the level of the liquid around the absorption cell falls slightly, *T* is lowered and liquid air is forced out of *C*, thus rapidly raising the level in the Dewar flask.

---

* Actually for the work described in the two following papers no thermocouple was used as we confined our attention to liquid nitrogen temperatures.

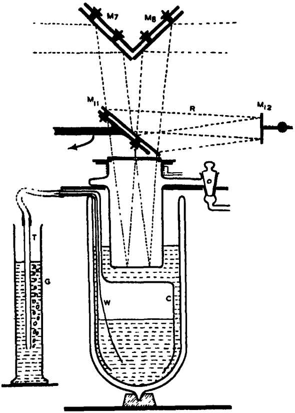


FIGURE 2. Absorption cell.

THE DETECTOR SYSTEM

We use a vacuum thermocouple of the Pfund type (Pfund 1937) specially constructed for us at the University of Michigan. The components used are bismuth/antimony and bismuth/tin alloys. The vacuum is maintained by a charcoal trap cooled in liquid air. The current from the thermocouple is taken through the shielded wires to a Kipp and Zonen Zc moving-coil galvanometer, the deflexions of which are amplified by an arrangement which combines the grid scheme mentioned by Pfund (Pfund 1929) with the photo-cell relay of Hill (Hill 1934), Jones (Jones 1934) and others. It is illustrated in figure 3.

The mirror of the primary galvanometer, whose deflexion we wish to measure, is M. An image of the lamp filament F is focused by lenses L_1, L_2 on to M and an image of the grid G_1 is produced on a second grid G_2. The bars and slots of G_1 are each 2 mm. wide and 4 cm. long and there are twelve slots altogether. G_2 is identical with G_1 except that the middle bar is made 4 mm. wide. If, when the image of G_1 is superposed on G_2 one-half of the spaces coincide, so that light passes through one-half of G_2, then the other half of G_2 will transmit no light. A biprism behind G_2 separates the light from the two halves and the lens L_3 focuses the two beams on to two photo-cells P_1, P_2, connected in series and in opposition. The output of these cells is fed into a galvanometer of moderate sensitivity (10^{-8} amp./mm. at 1 m.). The amplification produced can be conveniently

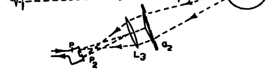


FIGURE 3. Photoelectric amplifier.

varied over a range wide enough for our purpose by altering the intensity of the light source by a variable resistance in series with it. We normally use an amplification of about 400, which would give a zero variation due to Brownian motion of the order of 5 mm. if the primary galvanometer were used critically damped. It has been found less arduous, however, to use the primary galvanometer rather over-damped by increasing the magnetic field. Although this increases the time of swing, there is a gain in both sensitivity and steadiness which amply compensates for this slight disadvantage. The linearity of the amplifier was carefully tested and found to be satisfactory as long as one was working inside the middle two-thirds of the total throw of the secondary galvanometer caused by one grid division deflexion of the primary galvanometer. The grid G_2 was mounted with a horizontal motion controlled by a fine pitch screw for adjustment of the 'zero position' of the secondary galvanometer spot so that deflexions always lay in this linear range. Because of the high sensitivity of this system to stray electrical and thermal disturbances great care was taken to shield it from such effects. In this connexion it was found to be extremely helpful to insert a window W over the slit S_3 at the entrance to the thermocouple housing. When working in the 3–4μ region the shutter consisted of a very thin piece of fused silica while for longer wave-lengths thin fluorite or rock-salt plates were employed. By this means all stray radiation of wave-length longer than that passed by the shutter was excluded from the thermocouple. Since the maximum of ordinary temperature radiation is in the neighbourhood of 10μ this was most effective for wave-lengths up to about 8μ, but even for longer wave-lengths it was helpful, particularly in excluding draughts. All of the apparatus (except the absorption cell) was enclosed in celotex boarding.

CALIBRATION

The instrument was calibrated by plotting the 3·4μ fundamental absorption band of hydrochloric acid gas. This has been very carefully recorded by Levin and Meyer (Levin and Meyer 1929) and affords the nearest approach to an absolute wave-length standard for this region of the infra-red. Three different gratings had to be calibrated, viz. a grating with 7200 lines to the inch used for wave-lengths of 3·6μ and under, a grating with 4800 lines to the inch used between 3·5 and 6μ and a grating with 1200 lines to the inch used for all wave-lengths greater than 6μ. For the two finer gratings the absorption of the HCl was directly plotted using a sufficiently narrow slit to separate the isotopic doublets. The values found for the grating constant k in the equation $\lambda = k \sin \theta$ were $10{\cdot}1832 \pm 0{\cdot}0005$ and $6{\cdot}7900 \pm 0{\cdot}0004$ for the 4800 and the 7200 gratings respectively. It will be noticed that the relation

$$k_{4800} = \tfrac{3}{2}k_{7200},$$

is not exactly fulfilled. This is due to a slight curvature of the 7200 grating which necessitated a small adjustment in the position of the mirror M_5 when this grating was inserted. The value of k for the 1200 grating was found from the relation

$$k_{1200} = 4k_{4800}.$$

It should be added that in the first equation λ is given in μ and θ is the angle between the zero position of the grating (corresponding to specular reflexion) and the position which puts the desired wave-length λ on the exit slit. With regard to the performance of the instrument it may be stated that slit widths corresponding to about 1 cm.$^{-1}$ can be used everywhere except in the regions of very strong atmospheric absorption.

This meant working with galvanometer deflexions of the order of 30 cm. in which the uncertainty was not more than 0·5 cm. In certain favourable regions slit widths of 0·5 cm $^{-1}$ have been used but this may be regarded as about the limit of the present set up.

The construction of this instrument would never have been possible but for the benevolent interest of two institutions, the assistance of which one of us (G.B.B.M.S.) now wishes to acknowledge very warmly. A Government Grant of the Royal Society provided for three of the four gratings. The Rouse Ball Foundation of Trinity College provided for the mirrors, the fourth grating and several indispensable plates of rock salt and sylvine necessary for the absorption cells. One of us (G.K.T.C.) is indebted to the Department of Scientific and Industrial Research and to Sidney Sussex College for maintenance grants.

REFERENCES

Barnes, R. B. 1930 *Phys. Rev.* **36**, 296.
Conn, G. K. T. and Sutherland, G. B. B. M. 1939 *Proc. Roy. Soc.* A, **172**, 172.
Hill, A. V. 1934 *J. Sci. Instrum.* **2**, 246, 281.
Jones, R. V. 1934 *J. Sci. Instrum.* **2**, 302.
Lee, E. and Wu, Cheng-Kai 1939 *Trans. Faraday Soc.* **35**, 1365.
Levin, A. and Meyer, C. F. 1929 *Phys. Rev.* **34**, 44.
Pfund, A. H. 1927 *J. Opt. Soc. Amer.* **14**, 337.
—— 1929 *Science*, **69**, 71.
—— 1937 *Rev. Sci. Instrum.* **8**, 417.

B. RAMAN SPECTRA

By G. B. B. M. Sutherland and Cheng-Kai Wu

The special Dewar vessel used (figure 4) is a modification of a type introduced by De Hemptinne (De Hemptinne and others 1938) for the study of liquids at low temperatures, in which we have turned the scattering column into the vertical position. This makes the apparatus just as suitable for solids as for liquids and more suitable than De Hemptinne's when only a small quantity of the substance under examination is available. The refrigerant is put in the outer tube *B* and the substance under examination is introduced directly into the inner tube *A*. The temperature is controlled both by the temperature of the refrigerant chosen and also by the thermal contact between the refrigerant and the substance in *A*. In the latter method the solid is put into a still narrower observation tube inside *A* and thermal contact made with the refrigerant in *B* by having a suitable amount of petrol ether in *A*. Only the upper part of the Dewar is silvered. The narrow end is jacketed and by means of a simple self-circulating system is kept from being warmed by the heat from the mercury arc. The liquid which was circulated by convection was a saturated solution of cobalt sulphate in water. This is very effective in absorbing all radiation of wave-length greater than 4360 A and in cutting down the continuous background which militates so much against the observation of very weak lines.

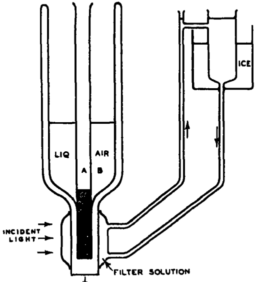


FIGURE 4. Apparatus for obtaining Raman spectra at low temperatures.

In the work reported in the following paper on HCl we actually kept everything at liquid nitrogen temperature. The HCl was condensed first as a liquid and subsequently solidified. This was found to be necessary in order to get a satisfactory mass of scattering material. To be certain of its being always at liquid nitrogen temperatures, the HCl was contained in a small observation tube which was put inside *A*, the latter also being well filled with liquid nitrogen. The scattered light was reflected into the spectrograph by means of a right-angle prism. Several stops were used to ensure that the minimum amount of reflected light entered the slit. The spectrograph employed was the Hilger E 439 instrument having an aperture *f*/3. The plates were Agfa Isopan I.S.S. which were rapid, free from halation and sufficiently fine-grained. The mercury arc used was a Hanovia 150 W lamp with oxide coated electrodes. It gave a few weak lines in addition to the mercury lines but these were not troublesome. It was placed at one focus of an elliptical mirror and the Raman observation tube at the other. The wave-lengths of the Raman lines were determined by a Hartmann interpolation formula and also checked against standard iron arc lines. The frequencies recorded cannot be in error by more than 5 cm.$^{-1}$.

REFERENCE

Hemptinne, M. de, Jungers, J. and Delfosse, J. M. 1938 *J. Chem. Phys.* **6**, 319.

Investigations on the vibration spectra of certain condensed gases at the temperature of liquid nitrogen

II. The infra-red and Raman spectra of hydro- and deuterochloric acid at liquid nitrogen temperatures

By E. Lee, G. B. B. M. Sutherland and Cheng-Kai Wu

(*Communicated by R. G. W. Norrish, F.R.S.—Received* 16 *May* 1940)

The infra-red absorption spectrum of solid hydrochloric acid has been obtained at the temperature of liquid nitrogen (i.e. approximately 20° below its transition point), in the region of the fundamental frequency. The results obtained agree in general with the earlier observations of Hettner and disagree with those of Shearin who reported discrete rotation lines. The doublet band observed by Hettner has been resolved into two separate bands with maxima at 2701 and 2744 cm.$^{-1}$ and indications of a third band were obtained. Some fine structure very different from that reported by Shearin was observed on each of the two bands which were unsymmetrical in shape, being steeper on the low frequency side. The Raman spectrum was obtained at the same temperature and also showed a doublet structure in contrast to earlier observations at higher temperatures which had given only a single line. The maxima of the Raman lines lay at 2709 and 2759 cm.$^{-1}$. The explanation of this doubling is to be sought in the manner in which the hydrochloric acid molecules associate in the solid state, since the alteration in the fundamental frequency from that found in the gaseous state shows that strong interaction must exist. The absorption due to the fundamental frequency of deuterochloric acid was observed under the same conditions.

Here three maxima were obtained at 1965, 1982 and 1992 cm.$^{-1}$ and indications of more structure. These results show that the association of HCl and DCl molecules in the crystal below the transition temperature can probably take place in several different ways. Such a conclusion is in agreement with the X-ray investigations which revealed a symmetry so low that the crystalline structure could not be determined.

It has long been known that at 98° K crystalline hydrochloric acid exhibits a transition point, but what exactly takes place in this transition has never been settled. Above 98° K the crystal possesses cubic (close-packing) symmetry but on cooling below this temperature the symmetry alters to a very much lower order (Simon and Simon 1924): simultaneously there is a very sudden drop in the dielectric constant (Cone, Dennison and Kemp 1931, Smyth and Hitchcock 1933) and in the specific heat (Eucken and Karwat 1927). The density shows a slight increase over the same temperature range. One explanation which has been proposed to account for all these phenomena is that above 98° the molecules are free to rotate about their mean positions while below that temperature the HCl dipoles may be regarded as fixed in the lattice. One of the most direct means of obtaining evidence of such rotation would appear to be an investigation of the vibration-rotation spectra of HCl above and below the transition point. Such an investigation was first made by Hettner (1932, 1934) who obtained some surprising results from which however no definite conclusion could be reached. A later investigation by Shearin (1935), confined to the absorption spectrum below the transition point, gave results in disagreement with those of Hettner. From Shearin's results it appeared definite that rotation was taking place below the transition temperature. A simultaneous investigation on the Raman spectrum at temperatures just below the melting point (i.e. considerably above the transition point) by Callihan and Salant (1934) yielded results in substantial agreement with those of Hettner for that region. It was with the object of clearing up some of the apparent contradictions in the above work, and of obtaining more clear-cut evidence regarding the hypothesis of molecular rotation in the solid that the present investigation was undertaken.

INFRA-RED SPECTRUM OF HCl

The spectrometer and absorption cell have already been described in the preceding paper. It is necessary however to record certain experimental details as the spectrum obtained was not always reproducible in detail in spite of every attempt to make the conditions of observation exactly the same in each case. This does not simply mean that there were small irregularities in the plot of the spectrum, such as occur from experimental error in all infra-red work, but that a characteristic point, such as the position of maximum absorption, varied between small but definite limits.

The hydrogen chloride was prepared by the action of sulphuric acid on ammonium chloride. It was dried by passing through concentrated sulphuric acid, was condensed by liquid nitrogen and re-evaporated, only the middle portion being collected. A peculiar phenomenon which had been noticed previously by Giauque and Wiebe (1928), was also observed by us, viz. the occurrence of a pink form of the normally white solid. When the acid was condensed *in vacuo* on to a glass surface previously cooled to the temperature of liquid nitrogen, a bright reddish pink coloration was obtained. This colour disappeared quickly on allowing it to warm up and re-cooling this white form did not bring back the colour. If on the other hand a small reservoir containing HCl was immersed in liquid nitrogen so that it condensed more slowly the coloured form was hardly ever obtained. The intensity of the colour varied from a very faint pink to a very deep red but in an unpredictable manner. No serious attempt was made to investigate this phenomenon as the absorption spectrum did not exhibit corresponding vagaries. It might be added, however, that Giauque and Wiebe were unable to get rid of it using drastic purification processes, and that we still found it in some later work in which the HCl was prepared by the action of distilled water on phosphorous trichloride of B.D.H preparation.

The HCl gas was admitted to the absorption cell after the base of the latter had been immersed in liquid nitrogen for some 10 or 15 min. The correct amount to admit to give a layer of suitable thickness was found by trial and error. If we assume that the HCl formed a uniform layer on the mirror then the thickness of the layer would have been approximately $0{\cdot}5\mu$. Since some of the HCl certainly condensed on the sides of the cell as well as on the base it might seem reasonable to take the above figure as an upper limit to the thickness. This is not a safe assumption however, as later experiments on germane indicated that the layer was thicker in the centre of the mirror than round the edges.

In making the first plots of the spectrum we did not find the percentage absorption at each point but simply recorded a total energy curve (figure 1). This was possible because the region in which we were working is practically free from atmospheric absorption. When once a plot had been made of this energy background it was therefore unnecessary to repeat it each time provided its intensity was determined at one or two points in the course of the run. In practice, this saving of work proved indispensable if we wished to cover the whole absorption spectrum in one run under high dispersion. It was important also to get the spectrum plotted as quickly as possible in order to be able to verify that changes were not taking place in the spectrum of the layer with time, and to make other such tests before the liquid air had to be replenished, as the latter could not be done without disturbing conditions. The slit widths employed are indicated on the diagrams and were generally of the order 1·5 cm.$^{-1}$. As the lines observed by Shearin were about 10 cm.$^{-1}$ apart we should have had no trouble whatever in picking them up. Our resolving power was of course much higher than that available to Hettner who employed a simple fluorite prism.

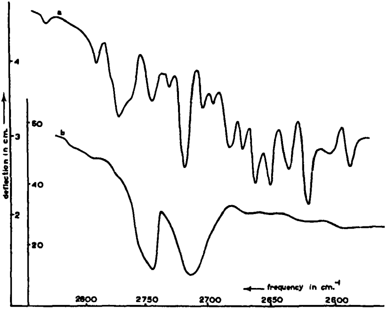


FIGURE 1. Absorption of HCl at liquid nitrogen temperatures. Curve *a* by Shearin; curve *b* by authors.

The general form of the absorption curve found by us is shown in figure 1 (curve *b*) where it is contrasted with that found by Shearin (curve *a*). It will be seen that it differs entirely from Shearin's but closely resembles that found by Hettner which is reproduced in figure 2. A possible explanation of the curious curve obtained by Shearin is that the variations found by him are due, not to HCl, but to water vapour in the atmosphere. This suggestion may seem contrary to our earlier remark that there is no appreciable atmospheric absorption in this region. It must be remembered however that Shearin was using the fourth order of a grating with 800 lines to the inch and so if separation of orders were incomplete he would have superimposed on the 2700 cm.$^{-1}$ radiation the fifth order of 3340 cm.$^{-1}$ radiation. The latter radiation is very strongly

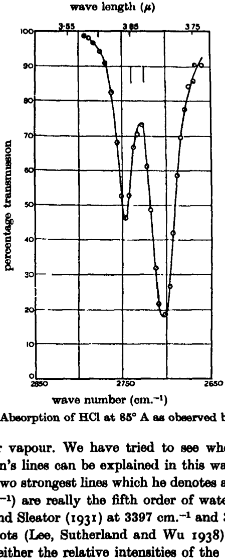


FIGURE 2. Absorption of HCl at 85° A as observed by Hettner.

absorbed by water vapour. We have tried to see whether some of the strongest of Shearin's lines can be explained in this way. It seems quite probable that the two strongest lines which he denotes as +4 (2718 cm.$^{-1}$) and −4 (2622 cm.$^{-1}$) are really the fifth order of water vapour lines recorded by Plyler and Sleator (1931) at 3397 cm.$^{-1}$ and 3277 cm.$^{-1}$.

In our earlier plots (Lee, Sutherland and Wu 1938) we were troubled by the fact that neither the relative intensities of the two peaks nor the exact positions of their maxima were quite reproducible. The former was finally traced down to having too thick a layer and not having entirely compensated for the effects of the short-wave radiation. The combination of these two factors meant that there was complete absorption for one peak when it seemed that the absorption was only about 85 % complete. The latter trouble has not yet been completely cleared up but it seems to have been largely due to variations in intensity in the short-wave radiation which again was not completely compensated for in the earlier experiments. It was only in the later experiments that the compensating shutter (which reflected the same amount of short-wave radiation as the filter shutter) was introduced. Also in the later experiments the background was plotted point by point simultaneously with the absorption. The typical curve obtained from those later experiments is shown in figure 3. Here it will be noticed that the two peaks are completely separated and are very similar in appearance being very steep on the long wave side and falling off more slowly on the short wave side where indeed they show definite signs of structure. Very weak maxima and minima were also observed at both ends of the main region of absorption and on two occasions a quite definite minor peak was found at 2777 cm.$^{-1}$. It may be remarked that Hettner observed similar weak maxima at the long wave end of the absorption region. In fact our results generally substantiate Hettner's and our higher resolving power has enabled us to show that indications of a fine structure do exist but it is entirely different from that reported by

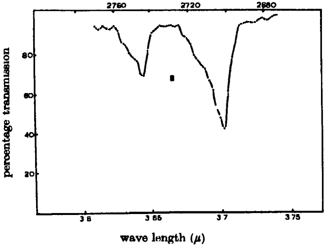


FIGURE 3. Absorption of HCl at liquid nitrogen temperatures under high resolving power.

Shearin. Regarding a slight variation in the position of maximum absorption of the lower frequency peak (between 2704 and 2714 cm.$^{-1}$) observed in the earlier runs it is interesting that there is a variation in the position of this absorption maximum in the two papers of Hettner. The first paper shows it very close to 2715 cm.$^{-1}$ while the later paper gives it at 2708 cm.$^{-1}$. We tried the effect of condensing the HCl at a temperature considerably above that of liquid nitrogen and then subsequently cooling to that temperature but the resulting curve was the same as that obtained by immediate condensation on a surface at liquid nitrogen temperatures.

INFRA-RED SPECTRUM OF DCl

The spectrum of crystalline deuterochloric acid was next investigated under the same conditions as it was hoped that the explanation of the two absorption peaks might be helped forward if one knew how they were affected by the substitution of deuterium for hydrogen. The DCl was prepared from D_2O obtained from the Nordiska Co. (purity 99·6 %) and PCl_3 of B.D.H. preparation, and purified in the same way as the HCl. We might remark that the pink form was again occasionally obtained but the coloration was much fainter.

The region of the absorption was now not free from water vapour lines. As before we plotted the energy background simultaneously and calculated percentage absorption for each individual point. The results are shown in figure 4 in which we have also included the strongest water vapour lines of the background to make it clear that certain kink in the DCl curve are probably due to their presence. The general form of the absorption is markedly different from that of HCl in that it exhibits a third absorption maximum This was carefully checked and there can be no doubt about its reality. The temperature of deposition of the layer was that of liquid nitrogen.

RAMAN SPECTRUM OF HCl

The preparation and purification of the HCl were exactly as for the absorption spectrum just described. The gas was first condensed as a liquid, which was subsequently frozen and cooled to the temperature of liquid nitrogen where it was held during the whole of the exposure. The apparatus has been described in the preceding paper. Two Raman lines were observed, the frequency shifts being 2709 ± 5 and 2759 ± 5 cm.$^{-1}$. The lines were not sharp, each one being about 15 cm.$^{-1}$ broad, and the high frequency one was the weaker of the two. This corresponds to the intensity relationship found in absorption. Figure 5 is a reproduction of the microphotometer trace. This is the first time that two Raman frequencies have been obtained from solid HCl. The earlier work of Callihan and Salant (Callihan and Salant 1934) showed a single broad line with a maximum at 2764 cm.$^{-1}$. This was done, however, just below the melting-point (about 155° K) and so a long way above the transition point. It agrees very well with the single absorption maximum found by Hettner (Hettner 1934) at 2768 cm.$^{-1}$ for HCl above the transition point. The

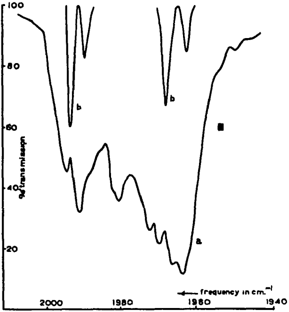


FIGURE 4. Absorption of DCl at liquid nitrogen temperatures. Curve *a* due to DCl; curve *b* shows strongest water vapour lines in atmospheric absorption at this region.

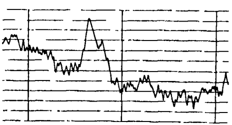


FIGURE 5. Raman spectrum of HCl at liquid nitrogen temperatures showing doublet structure.

Raman spectrum of the liquid also shows only a single rather broad frequency with its maximum at 2770 cm.$^{-1}$ (Salant and Sandow 1931) in spite of a recent attempt to search for a second frequency (Vodar, Freymann and Ta 1938).

DISCUSSION

The general substantiation of Hettner's observations as opposed to those of Shearin is the first important result of this investigation. Even if the structure reported by Shearin had been confirmed, it could hardly have been interpreted as due to rotation of the molecule in view of the recent work of West (1939) who has shown that all trace of rotational structure disappears for the gaseous spectrum at high pressures. Moreover no rotational structure has been observed in the liquid nor in the solid above the transition point (West 1939; Hettner 1932, 1934) and below the transition point is the least likely region of all for free rotation of the molecules to take place. We have remarked that in our curves there have been indications of fine structure. Whatever this structure may be due to (and we shall discuss its interpretation later) it cannot be a rotational fine structure. We shall consider first what light our work throws on the gross structure of this absorption band.

The most puzzling feature is the doublet structure as compared with the single absorption maximum found above the transition temperature and in the liquid. We have now resolved the doublet into two quite distinct bands and we have found indications of a weak third maximum at 2777 cm.$^{-1}$. We have also found that the corresponding absorption in solid DCl consists of three maxima. These facts when taken with Hettner's observation that at temperatures below 37° K a third peak appears at 2777 cm.$^{-1}$ (which is of greater intensity than the other two) make it appear likely that there are three separate maxima of absorption in solid HCl below the transition point and that the relative intensity of the three absorptions depends on the temperature of observation and possibly on the conditions of deposition. The question as to why there should be more than one maximum of absorption has been discussed by Hettner. He has shown conclusively that the doubling cannot be the unresolved contour of a vibration-rotation band and that it is extremely unlikely that it is due to combination of the intramolecular HCl frequency with an intermolecular lattice frequency as the magnitude of the latter fails to account for the specific heat curve. The isotope effect provides additional evidence against the hypothesis of a lattice frequency, as the latter would not be expected to alter very much in substituting deuterium for hydrogen in a molecular lattice. This would mean that the doublet separation would have remained practically unchanged in going from HCl to DCl, whereas it was found to alter almost exactly by a factor of $\sqrt{2}$. This is just what one would obtain from the individual alteration of each frequency by that factor.

The recent work on the effect of association between molecules on the intramolecular frequencies of the individual molecules is very relevant to this problem. It has been found by several workers that association between molecules (such as water, certain alcohols and acids) containing an OH group alters the frequency of that group by roughly 200–400 cm.$^{-1}$ depending on the strength of the so-called "hydrogen bond" thus formed. Now this is exactly the same order of magnitude as the change in frequency from gaseous HCl (2889 cm.$^{-1}$) to solid HCl (2701 and 2744 cm.$^{-1}$). There must, therefore, be a very strong interaction between the neighbouring HCl molecules in the solid and even in the liquid state (2785 cm.$^{-1}$). The simplest explanation of the multiple absorption maxima is that the HCl molecules may 'associate' in more than one way. The presence of the two well established maxima would show that there are two main modes of association but the presence of the third peak on certain occasions shows that a third structure is possible though presumably much less stable than the other two, except at very low temperatures. This interpretation would also explain why the X-ray studies of HCl below the transition point revealed such a very low symmetry, whereas above the transition point where the X-ray studies give a normal cubic structure, the absorption found by Hettner had a single peak. In this connexion we might mention the work of Vodar and others (Vodar *et al.* 1938) on the infra-red spectrum of liquid HCl in the region of the second overtone. They observed in addition to the normal band at 8130 cm.$^{-1}$ a much weaker band at 7752 cm.$^{-1}$. They have suggested that the former absorption corresponds to free HCl molecules in the liquid while the latter corresponds to associated molecules. This additional band reported by Vodar *et al.* has not been found in the neighbourhood of the fundamental nor of the first overtone and may therefore be spurious. If it is real, then we would modify the explanation given by those authors and suggest that the two bands in the liquid correspond to two different modes of association already taking place in the liquid We do not believe that such a thing as a 'free' HCl molecule exists in the liquid in view of the large shift in frequency which occurs in going from gas to liquid.

As regards the detailed structure of the absorption spectrum which we observed, it is clear from figure 3 that the low frequency side of each maximum is considerably steeper than the high frequency side. Although it is on the latter that the main signs of a fine structure appear there are also a few weak humps at each high frequency extremity. These seem to be quite real and have also been noted by Hettner, though he only reports them on the long wave extremity of the long wave maximum. Hettner has suggested that these are due to an interference effect but this seems very improbable in view of the fact that the thickness of the layer (about $0{\cdot}5\,\mu$) is only a fraction of the wave-length of the incident radiation ($3{\cdot}7\,\mu$) and that we have observed them at both ends of the absorption region. It seems much more probable that they are due to combination with very low lattice frequencies of the main intramolecular frequency. For such combination bands at low temperatures one would expect the summation bands on the high-frequency side would be much more intense than the difference bands on the low-frequency side. This would account for the unsymmetrical shape of each absorption peak.

A tabular summary of our results is given below together with those of earlier workers. It will be noticed that our values are slightly lower than Hettner's. Since we used a grating instrument of considerably higher resolving power we consider ours to be the more accurate values. As regards the Raman frequencies, it will be noted that there is a marked discrepancy between them and the infra-red frequencies both for the solid in the *B* form (i.e. below the transition point) and for the liquid. It seems larger than can be accounted for by experimental error, but no explanation can be offered for it at this stage. In the last two columns we

Infra-red and Raman spectra of HCl and DCl

	Type of spectrum	Gaseous state	Liquid state	Solid state *A*	Solid state *B*	$\Delta\nu$	$\Delta\nu/\nu$
HCl	Infra-red (earlier)	2889	2785	2768	2747	—	—
					2708		
	Infra-red (present)	—	—	—	2744	145	0·05
					2701	188	0·06
	Raman (earlier)	2886	2770	2764	—	—	—
	Raman (present)	—	—	—	2759	—	—
					2709		
DCl	Infra-red	2091	—	—	1992	99	0·05
					1982	109	0·05
					1965	126	0·06

have computed the absolute and relative changes in frequency in going from the gaseous state to the solid below the transition point. It is interesting to note that the percentage change in frequency is the same for HCl and DCl in view of the proposal (Badger and Bauer 1937; Fox and Martin 1937) that this quantity is directly proportional to the interaction energy in cases of association through a so-called 'hydrogen bond'.

This work was made possible by financial assistance from the Department of Scientific and Industrial Research and from the Royal Society. This assistance is now gratefully acknowledged by the author (G. B. B. M. S.) to whom it was made.

REFERENCES

Badger, R. M. and Bauer, S. H. 1937 *J. Chem. Phys.* **5**, 839.

Callihan, D. and Salant, E. O. 1934 *J. Chem. Phys.* **2**, 317.

Cone, R. M., Dennison, G. H. and Kemp, J. D. 1931 *J. Amer. Chem. Soc.* **53**, 1278

Eucken, A. and Karwat, E. 1927 *Z. Phys. Chem.* **112**, 467.

Fox, J. J. and Martin, A. E. 1937 *Proc. Roy. Soc.* A, **162**, 419.

Giauque, W. F. and Wiebe, R. 1928 *J. Amer. Chem. Soc.* **50**, 101.

Hettner, G. 1932 *Z. Phys.* **78**, 141.

— 1934 *Z. Phys.* **89**, 234.

Lee, E., Sutherland, G. B. B. M. and Wu, C. K. 1938 *Nature, Lond.*, **142**, 669.

Plylor, E. K. and Sleator, W. W. 1931 *Phys. Rev.* **37**, 1493.

Salant, E. O. and Sandow, A. 1931 *Phys. Rev.* **37**, 373.

Shearin, P. 1935 *Phys. Rev* **48**, 299.

Simon, F. and Simon, C. von 1924 *Z. Phys.* **21**, 168.

Smyth, C. P. and Hitchcock, C. S. 1933 *J. Amer. Chem. Soc.* **55**, 1830.

Vodar, B., Freymann, R. and Yoou Ta 1938 *J. Phys. Radium*, **9**, 282.

West, W. 1939 *J. Chem. Phys.* **7**, 795.

The flame spectrum of carbon monoxide

By A. G. Gaydon, Ph.D., A.Inst.P., *Radiation Gas Research Fellow, Chemical Technology Department, Imperial College, London, S.W. 7*

(*Communicated by A. C. G. Egerton, Sec.R.S.—Received* 13 *June* 1940)

[Plate 14]

The spectrum of the flame of carbon monoxide burning in air and in oxygen at reduced pressure has been photographed on plates of high contrast which display the band spectrum clearly above the continuous background. Greater detail has been obtained than has been recorded previously and new measurements are given.

The structure of the spectrum has been studied systematically. It is shown that the bands occur in pairs with a separation of about 60 cm.$^{-1}$, this separation being due probably to the rotational structure. Various wave-number differences are found to occur frequently, and many of the strong bands are arranged in arrays using intervals of 565 and 2065 cm.$^{-1}$.

The possible origin of the spectrum is discussed. The choice of emitter is limited to a polyatomic oxide of carbon, of which carbon dioxide is the most likely. The spectrum of the suboxide C_3O_2 shows some resemblance to the flame bands, but this molecule is improbable as the emitter on other grounds. A peroxide CO_3 is also a possibility, but no evidence for the presence of this has been obtained from experiments on the slow combustion of carbon monoxide.

Carbon dioxide in gaseous or liquid form is transparent through the visible and quartz ultra-violet, and the flame bands are not obtained from CO_2 in discharge tubes. Comparison with the Schumann-Runge bands of oxygen shows that it is possible that the flame bands may form part of the absorption band system of CO_2 which is known to exist below 1700 A if there is a big change in shape or size of the molecule in the two electronic states.

The electronic energy levels of CO_2 are discussed. Since normal CO_2 is not built up from normal CO and oxygen, an electronic rearrangement of the CO_2 must occur after the combustion process. Mulliken has suggested that the molecule in the first excited electronic state, corresponding to absorption below 1700 A, may have a triangular form.

The frequencies obtained from the flame bands are compared with the infra-red frequencies of CO_2 The 565 interval may be identified with the transverse vibration ν_2, indicating that the excited electronic state is probably triangular in shape. The 2065 interval cannot, however, be identified with the asymmetric vibration ν_3 with any certainty.

If the excited electronic state of CO_2 is triangular, then molecules formed during the combustion by transitions from this level to the ground state may be "vibrationally activated". This is probably the reason for many of the peculiarities of the combustion of carbon monoxide.

INTRODUCTION

The spectrum of the flame of carbon monoxide burning in air or oxygen is known to consist of a faint banded spectrum superposed on a strong continuous background which extends through the visible and near ultraviolet, being strongest in the region 4500 to 3500 A. Unless great care is taken to dry the gases, the OH band at 3064 A is invariably present in the spectrum of the flame. The spectrum has been studied by Weston (1925) and Kondratjew (1930), who have measured some of the strongest bands and attempted to arrange them into a fragmentary analysis. The same spectrum has also been observed by Fowler and Gaydon (1933) and Gaydon (1934*a*) in the afterglow of carbon dioxide in a discharge tube.

Weston expressed the opinion that the band system was produced by the direct interaction of carbon monoxide with oxygen, since it was present in the spectra of flames of pure CO and O_2, without the intervention of steam, but gave no definite opinion on the nature of the emitting molecule Kondratjew, and Fowler and Gaydon, attributed the bands to the CO_2 molecule, but this conclusion was apparently not entirely accepted by Bone (1932).

In all work hitherto the bands of the spectrum have been very faint, relative to the continuous background, and in the various published spectrograms, e.g. Weston (1925), Kondratjew (1930) and Bonhoeffer (1936), they are hardly visible, and measurements of such weak diffuse bands can hardly be expected to lead to any definite conclusions. The spectrum is indeed very complex and good measurable structure is difficult to get, but in the present investigation improved spectrograms have been obtained.

New measurements of the spectrum are presented and the structure of the band system and its origin are dealt with in the light of the more recent developments in our knowledge of molecular structure. The bearing of the knowledge gained from the spectrum to the problem of the combustion and after-burning of carbon monoxide will be dealt with in a later paper.

EXPERIMENTAL

The spectrum of the flame of carbon monoxide has been photographed with a Hilger E. 2 instrument. The carbon monoxide was stored in a cylinder and, in order to eliminate iron lines due to the presence of a trace of iron carbonyl formed in the cylinder, the gas was passed through a strongly heated tube packed with porcelain. No other precautions were necessary to purify the gas as no lines or bands in the spectrum were attributable to impurities, and there was no advantage in repeating Weston's careful investigations on the occurrence of the spectrum. No especial precautions were taken against water vapour, and the OH bands were present on all plates, although when the gas was at high pressure (100 atm.) in the cylinder the gas was fairly dry and the OH bands only weakly present.

The chief difference between the present spectrograms and earlier ones is that while previous authors have used relatively fast photographic plates with low contrast, for the present investigations plates of the highest contrast have been used. Most of the work has been done with Ilford Thin Film Half Tone and Panchromatic Half Tone plates developed in caustic hydroquinone; a few Process and Rapid Process Panchromatic plates have been used, but these proved less satisfactory. The advantage gained by using such high contrast may be judged by comparison of the published spectrograms. The composite photograph shown does not, however, display fully the advantage gained, as, while it is possible to measure dense plates which are well up on the straight portion of the photographic $D-\log E$ curve by using a strong viewing light, only weak plates with photographic density corresponding to the toe of the $D-\log E$ curve are satisfactory for enlargement.

It is known that the band structure is favoured relative to the continuous background by burning the gas at reduced pressure, and Kondratjew used this method for getting clearer bands. Some plates have been taken with carbon monoxide burning in pure oxygen at a pressure of about a third of an atmosphere using Weston's apparatus (Bone and Townend 1927), but the advantage gained was small. To use much lower pressure would have reduced the brilliance of the flame, and exposures on plates of high contrast would be very long In addition to the plates taken of carbon monoxide burning in oxygen at reduced pressure, many plates were taken of carbon monoxide burning in air at atmospheric pressure. The gas was burnt at a flattened silica jet, and air was mixed, at a T-piece junction in the tube, with the carbon monoxide before the gas was burnt at the jet, thus increasing the brilliance of the flame. Exposures ranged from 20 min. to 3 hr.

It was thought that admixture of carbon dioxide with the monoxide before burning might affect the relative intensity of the bands and the continuous portion of the spectrum. Some plates were accordingly taken of the flame of mixtures of carbon monoxide and dioxide burning in air and in oxygen, but apart from a marked extinguishing action by the carbon dioxide no change was noted in either the character of the flame or its spectrum.

A few experiments have been made on the slow combustion of carbon monoxide. These will be referred to briefly in later sections.

MEASUREMENTS

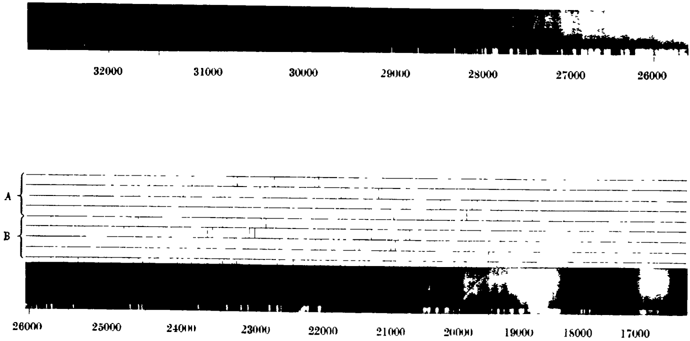
Twelve of the best plates have been measured carefully, iron arc comparison spectra being used. The means of these measurements are presented in table 1, which gives the wave-numbers *in vacuo* (in cm.$^{-1}$) of the maxima of the bands with intensities estimated visually on a scale of 1–5, and remarks on the character of the bands. The intensity estimates are to be regarded as intensity above the continuous background, which is very strong in the violet and falls off in the ultra-violet and at the visible end of the spectrum; thus clear bands in the red may be classified as intensity 5 although much longer exposures are required to record these bands than for bands in the violet classified as lower intensity. Many of the bands overlap, and such overlapping bands are grouped together in the table; bands which are probably double and are measured as such, but which are not clearly resolved, are bracketed.

Measurements of the bands agree among themselves to about 5 cm.$^{-1}$, and this probably represents the order of reliability of the measurements. The accuracy is, of course, limited by the diffuse nature of the bands, which do not in general show sharp heads or other features. The dispersion used was completely adequate. The present measurements cover the region 6250–3250 A, and agree reasonbly well with Fowler's and Kondratjew's measurements, which do not, however, extend beyond 5550 A, although Kondratjew's go further to the ultra-violet. The greater contrast has resulted in higher resolution and detail being obtained, and many bands previously measured as single are here shown to be more complex.

THE STRUCTURE OF THE BAND SPECTRUM

As will be seen from the plate the spectrum consists of a large number of narrow headless bands superposed on the continuous background. Many of the bands appear to be arranged in pairs, as has been remarked by Kondratjew. Some of the bands, especially those at the red end of the spectrum, show signs of being degraded towards the violet, while a few appear to be shaded in the other direction. This degradation of the bands is not, however, striking and may be due to overlapping of bands of dif-

OH 3064 A



The plate shows a composite enlargement made from spectrograms of the flame of carbon monoxide burning in oxygen. The spectrograms were taken on a Hilger E 2 quartz spectrograph using Half Tone and Panchromatic Half Tone plates. An iron arc comparison spectrum is shown and an approximate scale of wave-numbers *in vacuo* is given beside this. The bands given in the arrays in table 2 are indicated.

TABLE 1. WAVE-NUMBERS (*IN VACUO*) AND INTENSITIES OF THE BANDS OF THE CARBON MONOXIDE FLAME SPECTRUM

ν	I	ν	I	ν	I	ν	I	ν	I	ν	I
30765	2	27942	2	25715*	3	23505	3*r*	20870	4*r*	18456	2
732	2	921	2	660	2	464*	5	831*	5*v*	414	3
636	2*v*	844	2	601*	2	400	2	775	2*b*	385	2*v*
602	2*r*	769	1	561	3*d*	288	3	672	2*d*	222	3*b*
406	1	744	2	506*	3*v*	260	3	637	3*d*	162	2*v*
30271	3*v*	709	1	379	1	216*	3	583	3*d*	18070	2*v*
29868	1	644	4*v*	349	2*d*	077	5	556	2	17997*	3*v*
850	1	597	2*d*	320*	2*d*	23020*	5*v*	492	2	954	1*s*
821	1	533*	2*d*	288*	2*d*	22931	4	455*	5*b*	850	3*v*
714	1*b*	440	1*d*	237	1	876*	4	400	2*v*	790*	3*v*
668	2	292	1	189	4	843?	2	308	4*b*	681	2*d*
617	1*d*	272	1	143	1	720	2	259*	4*vs*	631	1
571	2	237	2	122	2	691	3	198	2*v*	577	1
542	2	098	1	097	2	653*	3*v*	092	4	545	1*s*
477	4*v*	069	2	070?	1	561	3	074	5	485	2
426	2*d*	27031	1	047?*	3	510	3	20022*	4*v*	430	1*v*
403	2*v*	26935	2	25003	3*r*	458	2*v*	19974	2*r*	347	2*r*
284	2*v*	852	3	24937*	3*v*	409	1	904	4*b*	288	3*vb*
247	1	824	2	774	3	359	3*r*	854	3*v*	213*	3*vb*
217	1	783	2*b*	713*	4	305	3*v*	797*	3*vs*	118	4
184	1	712	1	657	3	149	2*d*	707	2*d*	17046*	3
111	1	694	2	642?	3	118	3*d*	654	3*d*	16971	2
038	1	638*	3	613?	3	22096*	3*d*	616	3*v*	887	5
29002	2	511	1	582*	5*v*	21962	5	544	2*d*	865*	5*v*
28969	1	384	2	524*	3*d*	884*	5*v*	526	3*d*	735	5
699	2*b*	347	1	428	4*s*	796	4*v*	499*	3*d*	696	5*v*
664	2	301	2*d*	372*	4	740*	3	470	3*d*	607	4
609	2	272	2*d*	340	1*v*	677	2	403	2*s*	527	4*b*
574	3	231	1	253	1	644	2	349*	5*b*	458*	4*v*
538	1	218?	2	209	4	585	4	219	1	411	2
511*	3*b*	170	1	151*	4*v*	556	4	174	4	356	3
455	2	137	2	060	5	529*	5*v*	19124*	4*v*	326	3
360	1	119	1	24000*	3	439	3	18983	5	231	3*d*
277	3	082*	4*v*	23944	2	389*	3*v*	950*	5	190*	3*v*
243	3	26038*	2	893?	2*d*	326?	1	836	4	113	2
195?	2	25954	2	858?	2*d*	278	1*v*	782*	5	16036	1*v*
148	3*b*	932	2	815*	4*v*	208	3	721*	4		
087*	3*s*	903*	2	760?	1*v*	184	2	688	5*v*		
053	2*r*	897*	3	700	2*d*	146	2	611*	4*d*		
28028	1	845	3	634	3*d*	078	2	586	5*d*		
27997*	2*b*	792*	2*d*	617?	3	21022	3	557*	5*d*		
971	1	773	2*d*	595?*	3*v*	20974*	4*v*	530*	4*vd*		

b = broad, perhaps double; *d* = diffuse; *s* = sharp, *v* = shaded to violet; *r* = shaded to red.

ferent intensities producing this effect. It cannot be assumed for a polyatomic molecule, as it can for a diatomic molecule, that the direction of degradation of the bands will give a clue to the relative magnitudes of the initial and final vibrational frequencies, as the direction of degradation will of course depend on the various moments of inertia of the molecule in the two states and these will be dependent on the shape of the molecule as well as on the strength of binding which principally governs the vibrational frequencies.

In attempting a vibrational analysis, Fowler, from Weston's plates, used intervals of around 570 cm.$^{-1}$ and 370 cm.$^{-1}$, obtaining a number of fragmentary schemes with these frequency differences. Kondratjew arranged the bands in long series with intervals increasing from around 550 cm.$^{-1}$ to over 600 cm.$^{-1}$, this interval being provisionally identified with the interval of 667 cm.$^{-1}$ observed for the infra-red spectrum of CO_2. These long series, however, are not completely confirmed by the present measurements, the intervals and intensities showing rather disturbing irregularities in some parts. With such a very complex spectrum, consisting of diffuse bands every few angstroms and whose accuracy of measurement is so limited, great care must be taken not to force the analysis to fit any preconceived ideas. In order to avoid this difficulty as far as possible it was decided to examine the wave-number differences between all the strong bands systematically to find, without bias, those which occur most frequently.

The wave-number differences between all the strong bands, those listed as intensity 3 or greater in the wave-number range 29,000 to 24,700 cm.$^{-1}$ and 4 or greater in the range 24,700 to 16,000 cm $^{-1}$, were computed up to a difference of 2400 cm.$^{-1}$, this presumably covering the range of any vibrational frequencies likely to be encountered. The number of differences between 0 and 9, 10 and 19, 20 and 29 cm.$^{-1}$, etc. were then counted and the number plotted against the mean interval. This should show up at once any regularity in the spectrum, as wave-number differences corresponding to the various vibrational frequencies and simple combinations of these should show peaks on the curve. The result was admittedly disappointing, the curve showing little outstanding structure. There were fairly definite peaks (more than 15 differences in the 10 cm.$^{-1}$ interval) at 60, 1130 and 1500 cm.$^{-1}$, and possible peaks at around 155, 215, 365, 510, 565, 790, 910, 960, 1070, 1190, 1290, 1350, 1700, 1870, 1930, 2090 and 2260 cm.$^{-1}$. The 60 cm.$^{-1}$ interval was definite and corresponded to the obvious occurrence of the bands in pairs, as already noted.

In many cases the lower frequency component of these pairs is relatively sharp and often shaded to the violet, while the higher frequency component is rather stronger and broader, being possibly itself a close doublet in some cases. This structure is very probably due to the rotational structure of the band, and may correspond to the P and R or P and Q branches. It was decided to accept this as a working hypothesis and to re-examine the frequency differences using one component only; the lower frequency component was selected as being the sharper and often the more free from overlapping. The bands which appear as though they may be the lower frequency components of such doublets were selected by inspection and are marked in table 1 with asterisks. This selection was made before attempting further work on the analysis and the selection is apparently in error in a number of cases.

The grouping of the frequency differences obtained from these selected bands was plotted as shown in figure 1, and shows considerable regularity, the peaks being very much more definite than when using all the strong bands. In figure 1 the number of wave-number differences occurring in the intervals 0 to 19, 10 to 29, 20 to 39, 30 to 49, etc. are plotted against the mean intervals 10, 20, 30, etc. This plotting in 20 cm.$^{-1}$ groups which overlap does not of course give any more real information than by plotting in simple 10 cm.$^{-1}$ intervals, but as a smoother curve is obtained the regularities are more obvious.

As may be seen from figure 1, there are strong peaks (more than 15 differences in the 20 cm.$^{-1}$ interval) at 565, 1130 and 1500 cm.$^{-1}$, quite definite peaks at 345, 925, 1355, 1700, 2065 and 2260, and less certain peaks at around 160, 230, 515, 780, 1865 and 1915 cm^{-1}.

It is at once obvious that within the accuracy of the measurements,

$$\begin{aligned} 1130 &= 2 \times 565, \\ 1700 &= 3 \times 565, \\ 2260 &= 4 \times 565, \end{aligned}$$

and that

$$\begin{aligned} 2065 - 1500 &= 565, \\ 1500 - 925 &\doteq 565, \\ 1915 - 1355 &= 565, \\ 1355 - 780 &\doteq 565, \\ 780 - 230 &\doteq 565, \end{aligned}$$

and that

$$\begin{aligned} 2260 - 1915 &= 345, \\ 1700 - 1355 &= 345, \\ 1130 - 780 &= 345, \\ 565 - 230 &\doteq 345. \end{aligned}$$

It is therefore seen that all except three of the less marked differences may be built up from three fundamental differences, which will of course correspond to three vibrations of the molecule emitting the spectrum. In the choice of which are these three fundamentals we must, however, pass from the firm basis of observation to less certain speculations.

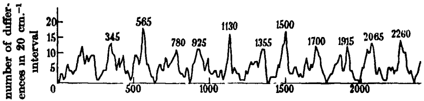


FIGURE 1

The 565 interval is a fairly obvious first choice as a fundamental, as this frequency and its multiples 1130, 1700 and 2260 are all so strong. The next choice might fall upon the 1500 cm.$^{-1}$ difference, which shows such a strong peak on the curve, but 2065 ($=1500+565$) or 925 ($\doteq 1500-565$) are possible It will be seen in the discussion which follows that the most likely emitter is carbon dioxide, and that the transition is to the ground state. If this is so, then it is very difficult to find a place for intervals of 1500 or 925, but that the 2065 interval could possibly be identified with the 2350 vibrational frequency of CO_2 if allowance is made for an anharmonic factor The choice of 2065 has also the advantage that the arrays presented in table 2 are then of compact form.

Using the intervals 565 and 2065, it is possible to build up two fairly

TABLE 2

A 1

		24582(5)*b*				20455(5)*b*	*2041*	18414(3)
		582				*551*		*564*
26082(4)	*2082*	24000(3)*b*	*2038*	21962(5)*b*	*2058*	19904(4)*b*	*2054*	17850(3)
576		*563*		*573*		*555*		*562*
25506(3)	*2069*	23437(3)*b*	*2048*	21389(3)*b*	*2040*	19349(5)	*2061*	17288(3)
569		*561*		*558*		*567*		*553*
24937(3)	*2061*	22876(4)	*2045*	20831(5)*b*	*2049*	18782(5)	*2047*	16735(5)
565		*571*		*572*		*560*		
24372(4)	*2067*	22305(3)	*2046*	20259(4)	*2037*	18222(3)		
557		*565*		*552*				
23815(4)	*2075*	21740(3)	*2033*	19707(2)				
555		*556*						
23260(3)	*2076*	21184(2)						

b = blended with a band used elsewhere in the arrays; *w* = weakly present on plate but not measured.

TABLE 2 (*continued*)

A. ii

		24642(3)				20556(2)		*b*
		582				*582*		
26137(2)	*2077*	24060(5)				19974(2)		*w*
576		*555*				*571*		
25561(3)	*2056*	23505(3)	*2066*	21439(3)	*2036*	19403(2)	*2056*	17347(2)
558		*574*		*569*		*567*		
25003(3)	*2072*	22931(4)	*2061*	20870(4)	*2034*	18836(4)		
575		*572*		*562*				
24428(4)	*2069*	22359(3)	*2051*	20308(4)		*w*		
570		*563*						
23858(2)	*2062*	21796(4)		*w*				
		w						

B i

		25288(2)	*2072*	23216(3)	*2070*	21146(2)		
		575		*563*		*563*		
26783(2)	*2070*	24713(4)	*2060*	22653(3)	*2070*	20583(3)	*2053*	18530(4)
565		*562*		*557*		*561*		
26218(2)	*2067*	24151(4)	*2055*	22096(3)	*2074*	20022(4)		
558		*556*		*567*		*552*		
25660(2)	*2065*	23595(3)	*2066*	21529(5)	*2059*	19470(3)		
563		*575*		*555*				
25097(2)	*2077*	23020(5)	*2046*	20974(4)				
573		*562*		*574*				
24524(3)	*2066*	22458(2)	*2058*	20400(2)				
		574		*546*				
w		21884(5)	*2030*	19854(3)				
		558						
23387(2)	*2061*	21326(1)		*w*				
		551						
w		20775(2)						

B. ii

		25349(2)	*2061*	23288(3)	*2080*	21208(3)		
		575		*568*		*571*		
26852(3)	*2078*	24774(3)	*2054*	22720(2)	*2083*	20637(3)	*2051*	18586(5)
580		*565*		*571*		*563*		
26272(2)	*2063*	24209(4)	*2060*	22149(2)	*2075*	20074(5)		
557		*575*		*564*		*548*		
25715(3)	*2081*	23634(3)	*2049*	21585(4)	*2059*	19526(3)		
572		*557*		*563*		*576*		
25143(1)	*2066*	23077(5)	*2055*	21022(3)	*2072*	18950(5)*b*		
561		*567*		*567*				
24582(5)	*2072*	22510(3)	*2055*	20455(5)*b*				
582		*548*		*551*				
24000(3)	*2038*	21962(5)	*2058*	19904(4)*b*				
563		*573*						
23437(3)*b*	*2048*	21389(3)*b*		*b*				
		558						
b		20831(5)*b*						

satisfactory arrays A. i and B. i, shown in table 2, and, using the doublet partners of the bands in these arrays, which lie about 60 cm.$^{-1}$ to the violet, two more arrays A. ii and B. ii may be formed. Nearly all the strong bands between 27,000 and 20,000 cm.$^{-1}$ are included in these four arrays. It is possible to force a few more bands into these arrays and to build up fragmentary arrays of a similar type using other bands, but these are unconvincing and do not give any more information about the structure of the spectrum.

The arrays are shown in table 2, the wave-numbers of the bands being printed in ordinary type and followed by the intensities in brackets, while the differences between the wave-numbers are printed in italics.

Subtraction of the wave-numbers of the bands in the second vertical row of B from the corresponding bands in the first row of array A gives differences of about 1355. Thus the two systems A and B, if displaced horizontally by one unit (2065 interval), are seen to be separated by the observed 1355 interval, and that a further displacement upwards or downwards by a number of 565 units will show the observed intervals 1920 (1355 + 565), 790 (1355 − 565), 230 (1355 − 2 × 565), 345 (3 × 565 − 1355). The fundamental corresponding to this displacement between the two arrays is uncertain. It would possibly correspond to a vibrational frequency of the molecule in the excited electronic state.

It must be remembered that it is possible to rewrite the arrays in table 2 either by shifts of one or more units between rows or by reflection of the array as a whole. The lay-out given here is chosen partly in order to bring the structure into line with what might be expected for carbon dioxide and partly because of the relative compactness as set up in this form. In a better defined spectrum, in which more accurate measurements could be made, the change of the intervals due to anharmonic factors, assisted by a study of the intensity distribution, would probably enable a more definite assignment of the fundamentals to be made, but this cannot be done here as the errors of measurement due to diffuseness and overlapping of bands produce deviations which are greater than the second-order differences caused by anharmonic factors.

The above discussion of the structure may be summarized by saying that the band system possesses a doublet structure with a separation of about 60 cm.$^{-1}$ and that there is very probably a vibrational frequency of about 565 cm.$^{-1}$ (increasing slightly for differences between bands at higher frequency) associated with the lower electronic state. Differences of 1500 or $2065 \pm n \times 565$ and $1355 \pm n \times 565$ occur, but it is uncertain which of these correspond to fundamental vibrations of the molecule.

THE ORIGIN OF THE BAND SPECTRUM

Weston has shown that the spectrum is emitted by pure dry carbon monoxide burning in oxygen. It is therefore certain that the emitting molecule can only contain the elements carbon and oxygen. It is equally certain that the spectrum is much too complex for a diatomic emitter, and in any case the various band spectra of oxygen O_2, carbon C_2 and carbon monoxide are well known and show no resemblance to the observed bands, so that these molecules may be ruled out. Kondratjew has discussed this question and given reasons for assuming that the bands are produced by carbon dioxide.

The band spectrum of ozone O_3 is well known in absorption, and Johnson (1924) has observed bands in the near ultra-violet in emission in a discharge tube source, but these are not similar to the carbon monoxide flame bands, and as there does not appear to be any particular reason for expecting the presence of ozone we may dismiss this from among the list of possible emitters.

The only other stable oxide of carbon is the suboxide C_3O_2. The absorption spectrum of this has been studied by Thompson and Healey (1936), they observed absorption in the region 2400–3300 A, the region between 2800 and 3300 being banded; these bands show a few strong pairs with a separation of about 100 cm.$^{-1}$ and a number of weaker pairs with separation 50–70 cm $^{-1}$, and Thompson and Healey use vibrational frequencies of 840, 2160, 1227, 540, 240 and 150 in attempting an analysis. The bands of course lie to rather shorter wave-lengths than the carbon monoxide flame bands, but in the region in which the systems overlap there is some agreement, especially with Kondratjew's measurements, which extend to shorter wave-lengths than those presented here. The infra-red absorption has been studied by Lord and Wright (1937), who find among the strongest frequencies one of 550 cm.$^{-1}$. Thus based entirely on spectroscopic evidence it would be possible to make out a good case for C_3O_2 as the emitter; the coincidence between the band measurements is not convincing but is as good as might be expected for comparison between the absorption and emission bands of a polyatomic molecule, and the doublet separation 50–70 cm.$^{-1}$ and the frequency difference 540 or 550, which might be identified with the observed 565 interval, are favourable to the identification. The chemical aspect is, however, less satisfactory; the production of a lower oxide seems unlikely and there is no positive evidence for its presence; the flame has not been observed to show the absorption spectrum of C_3O_2; the suboxide is supposed to decompose to $CO_2 + C_2$ and this C_2 should show spectroscopically by emission of the Swan bands, which are not observed. We may conclude that C_3O_2 is unlikely but not impossible as the emitter.

The existence of a carbon peroxide CO_3 has been presumed in certain proposed chain mechanisms for the combustion (e.g. Lewis and von Elbe 1938). There appears to be no experimental evidence for the existence of CO_3 or other oxides of carbon. Some preliminary experiments have been conducted on the slow combustion of carbon monoxide with a view to testing for peroxide formation. A slow flow (about 100 c.c. per min.) of a 3:1 mixture of oxygen and carbon monoxide (the gases being roughly dried by passage through long phosphorus pentoxide drying tubes) was passed through a quartz tube (length 20 cm., volume 250 c.c.) heated in a furnace to 595° C. The gases from the tube were passed direct through only about 10 cm. of quartz tube into an acid solution of potassium iodide. No iodine was liberated even after prolonged bubbling of the emergent gases through the solution, although tests with baryta water indicated that there was fairly rapid carbon dioxide formation under these conditions. The absorption spectrum of the gases in the quartz tube was also examined under these conditions, but no absorption at all was observed as far as 2100 A, the limit of the spectrograph used. This failure to detect the presence of a peroxide is not of course conclusive, as short-lived radicals like OH and CH and unstable compounds like peroxides may be present in very small quantities and not show absorption although they may give strong emission spectra Nevertheless, in view of the complete absence of experimental evidence for the existence of CO_3 or other peroxide, these compounds could only be regarded as a possible but very unlikely source of the flame bands.

CARBON DIOXIDE AS THE EMITTER

Carbon dioxide molecules are the most likely source of the banded part of the carbon monoxide flame spectrum. This opinion has already been expressed by Professor A. Fowler and the author (1933) and Kondratjew (1930) and Kondratjewa and Kondratjew (1937*a*, 1937*b*). There are, however, difficulties on the spectroscopic side in assigning the bands to this molecule. The absorption spectrum of CO_2 has been studied by Liefson (1926) and by Price and Simpson (1938*a*), who find that the gas is transparent all through the visible and quartz ultra-violet region and that the first absorption bands lie below 1700 A, and strong absorption is only observed still further to the ultra-violet. Eiseman and Harris (1932) have similarly shown that liquid carbon dioxide is also transparent at least to 2150 A.

Since electronic selection rules are not rigorously obeyed in the liquid state (cf. absorption by liquid oxygen), this completely rules out the possibility of there being any low-lying electronic levels of CO_2. In flame sources there is not sufficient energy to excite atoms or molecules to high electronic states; thus the CO flame spectrum itself does not show well-known band systems, such as the CO Third Positive or Angstrom bands or any oxygen emission which would require a fairly high excitation energy; it therefore follows that if CO_2 is the emitter the electronic transition definitely involves the ground state. Also the emission spectrum of carbon dioxide has been studied thoroughly in discharge tubes and, apart from the afterglow which is probably only a flame due to recombination at low pressure, no bands corresponding to the flame have been observed.

If the band spectrum is emitted by CO_2 then there appears to be only one way of reconciling these facts, that is by assuming that the molecule has a very different shape or size in the ground and excited electronic states. The possible effect of such a change of size is best seen by comparison with the Schumann-Runge band system of oxygen, which affords probably the best diatomic example of a big change of size (internuclear distance) of the molecule in the two electronic states. In this case the absorption bands observed by Schumann lie below 1900 A and require the electronic energy 49,845 cm.$^{-1}$ plus a large vibrational energy to account for the intense vibration of the molecules in the excited state set up by the sudden change in equilibrium internuclear distance when the electronic transition occurs. Thus only bands on one limb of the very open Franck-Condon parabola, those bands with high values of v', the upper vibrational quantum number, are observed in absorption. In emission in a high tension arc in oxygen Runge observed bands of this system in the region 4400–3100 A. These bands correspond to the other limb of the Franck-Condon parabola, the transitions involving the electronic energy change 49,845 cm.$^{-1}$ less a large amount of energy due to vibration of the molecules in the ground state. These Runge emission bands do not, however, appear in discharge tube sources, this being due presumably to the inability of that type of excitation, which is chiefly by electron impact, to excite the oxygen molecule to an electronic state in which the internuclear distance is so much greater. For a tri-atomic molecule the occurrence of such an extreme change is much more likely than for a diatomic molecule because, in addition to a change in internuclear distance, there is also the possibility of a change of shape. For the sulphur dioxide afterglow (Gaydon 1934*b*) a less extreme case of this difference in region between the emission and absorption spectrum is indeed observed; the absorption from the ground state lies around 3000 A, while the afterglow bands, which are known from the analysis to definitely involve the ground state, lie around 4300 A. Thus, if there is a big change in the size or shape of the CO_2 molecule, the big difference in the regions of absorption (< 1700 A) and emission (3000–5000 A) and the non-appearance of the bands in emission in discharge tubes might be accounted for. On the basis of the above discussion it will be assumed, at all events tentatively, that the carbon monoxide flame bands are due to CO_2 and correspond to transitions to rather high vibrational levels of the ground electronic state. The vibrational levels of the CO_2 molecule are known from the study of the infra-red absorption spectrum (Dennison 1932, 1933 and references given therein). The three fundamental vibrational frequencies are ν_1 (symmetrical) about 1300 cm.$^{-1}$ (this is actually split into two by degeneracy and resonance with ν_2), ν_2 (transverse) $= 667$ cm.$^{-1}$, and ν_3 (asymmetrical) $= 2350$ cm.$^{-1}$. Some of these frequencies might be expected to appear as differences between the wave-numbers of the carbon monoxide flame bands. Kondratjew (1930) identified the interval of about 600 cm.$^{-1}$ with ν_2 (667 cm.$^{-1}$) and the present attempt at analysis also shows up strongly this interval, which has a value of about 565 in the region of the spectrum studied here; it is of course to be expected that the interval will be less than 667 for transitions to high vibrational levels owing to the anharmonic factor, some of Kondratjew's intervals are between bands at higher frequency than those presented here, and this would account for his giving the average value of the interval as high as 600 cm $^{-1}$. The other observed intervals of the flame bands are less easily accounted for. If the 1500 interval were chosen as fundamental it could, on account of its frequent occurrence, only be assigned to the lower electronic state, and in this case could not be due to carbon dioxide. If, however, the 2065 interval is selected as the fundamental, then it may be possible to identify this with the asymmetric $\nu_3 = 2350$ of CO_2. This agreement cannot be considered really satisfactory as the discrepancy seems very large to be accounted for by anharmonic factors, but it is just possible that if the molecule is in a highly excited level of the transverse vibration the asymmetrical vibrational frequency may be lowered quite considerably. The symmetrical frequency of about 1300 cm.$^{-1}$ for CO_2 is not apparent from the analysis of the flame bands. Other weak frequency differences which are observed in the flame spectrum, such as 345 or 1355 cm.$^{-1}$, might be assigned to the upper electronic state.

THE ELECTRONIC LEVELS OF CARBON DIOXIDE

It is known that, since carbon dioxide is diamagnetic, the ground electronic state is of the $^1\Sigma$ type (Mulliken 1935, p. 721) and it has been pointed out by Herzberg (1932) that normal CO in its $^1\Sigma$ ground state and an unexcited oxygen atom O(^{3}P) cannot give this $^1\Sigma$ level. Herzberg assumed therefore that the products of dissociation of normal CO_2 would probably be not CO($^1\Sigma$)+O(^{3}P) but CO($^1\Sigma$)+O(^{1}D) This hypothesis is open to the objection that CO_2 is known to be a linear symmetrical molecule, and if it was built from CO($^1\Sigma$) and O(^{1}D) this would probably not be the case. Schmid (1936) has suggested that the normal CO_2 is formed from CO($a\ ^3\Pi$) and a normal ^{3}P oxygen atom. This $a\ ^3\Pi$ level of CO has an energy of about 6 eV above the ground $^1\Sigma$ state and is formed from a ^{5}S carbon atom and a normal ^{3}P oxygen atom, giving a symmetrical construction for the CO_2 molecule of O(^{3}P)+C(^{5}S)+O(^{3}P). The normal state of carbon is ^{3}P which is some 4·4 eV below the ^{5}S level. It has been recognized (e.g. Van Vleck (1934) and Voge (1936)) that quadrivalent carbon, as in hydrocarbons and saturated organic compounds generally, is in this excited ^{5}S state.

Since in the combustion we have no evidence for the prior excitation of the carbon monoxide to the $a\ ^3\Pi$ level, from which the Cameron band system of CO (bands at $\lambda\lambda$ 2258, 2369, 2491, etc.) would be emitted, it may be supposed that in the combustion process CO_2 is formed from the unexcited CO($^1\Sigma$) and oxygen and that the CO_2 so formed must then undergo an electronic transition to reach the ground state.

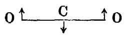
Mulliken (1935, p. 738) has also discussed the nature of the excited electronic states of CO_2 and has arrived at some very interesting conclusions; the first excited levels above the ground state are likely to be $^1\Pi$ or $^3\Pi$, but the transition to the ground state would be forbidden by the selection rules for linear symmetrical molecules; Mulliken then proceeds to identify, tentatively, the weak absorption centred around 1600 A with this forbidden transition, suggesting that the selection rules might be weakened if the equilibrium form of the molecule in the upper level is triangular.

DISCUSSION

It has been shown that the possible emitter of the flame bands can be limited to certain oxides of carbon, of which CO_2 is the most likely. The spectrum of the flame is not, however, the same as any known spectrum of CO_2 obtained in either emission or absorption. By assuming that the molecule in the excited electronic state is of different shape (probably triangular) from the linear unexcited molecule it may still be supposed that the bands of the flame are emitted by CO_2 in an electronic transition to the ground state.

The analysis does not prove as conclusively as might be expected that the lowest level of CO_2 is involved. The analysis of the spectrum and the interpretation thereof is, however, complicated by lack of knowledge of the vibrational selection rules for electronic transitions in polyatomic molecules. Herzberg and Teller (1933) and Ku (1933) have discussed these, but such rules as have been derived neglect important anharmonic factors, and cannot be applied without a knowledge of the type of the electronic transition and the symmetry of the molecule in both electronic states.

If the 565 wave-number difference observed in the flame bands is identified with the transverse vibration ν_2



of carbon dioxide, then the strong excitation of this form of vibration strongly supports the hypothesis that the excited state of CO_2 is triangular, as the transition to the linear state would obviously leave the molecule deformed from the linear shape in such a way that the transverse vibration would be strongly excited. If the 2065 interval in the flame bands is tentatively identified with the asymmetric ν_3, this might indicate that the excited molecule itself has an asymmetrical form

If the flame bands do, as seems likely from this discussion, correspond to an electronic transition from a triangular electronically excited CO_2 molecule to the linear unexcited state of the molecule, then the molecules after this transition will be left with a very large vibrational energy. The molecules formed would not therefore be in thermal equilibrium and might be regarded as 'vibrationally activated' if this vibrational energy required an appreciable time to be dissipated. This is probably the cause of many of the peculiarities of the combustion of carbon monoxide, with which it is hoped to deal more fully in later work.

It is a pleasure to express my thanks to Professor A. C. Egerton for his continued interest and many helpful suggestions. I also wish to thank Dr D. M. Newitt and Dr R. W. B. Pearse for helpful discussion, and Messrs Radiation, Ltd., for the Fellowship during the tenure of which these investigations have been made.

REFERENCES

Bone, W. A. 1932 *J. Chem. Ind.* p. 962.

Bone, W. A. and Townend, D. T. A. 1927 *Flame and combustion in gases.* Longmans.

Bonnhoeffer, K. F. 1936 *Z. Elektrochem.* **42**, 449.

Dennison, D. M. 1932 *Phys. Rev.* **41**, 304.

— 1933 *Phys. Rev.* **43**, 716.

Eiseman, B. J. and Harris, L. 1932 *J Amer. Chem Soc* **54**, 1782.

Fowler, A. and Gaydon, A. G. 1933 *Proc Roy. Soc.* A, **142**, 362

Gaydon, A. G. 1934*a* *Nature, Lond*, **133**, 984.

— 1934*b* *Proc. Roy. Soc.* A, **146**, 901.

Herzberg, G. 1932 *Z. phys. Chem* B, **17**, 68

Herzberg, G. and Teller, E. 1933 *Z. phys. Chem.* B, **21**, 410.

Johnson, R. C. 1924 *Proc. Roy Soc.* A, **105**, 683.

Kondratjew, V. 1930 *Z. Phys.* **63**, 322.

Kondratjewa, E. and Kondratjew, V. 1937*a* *Act. Physicochim. U.S S.R.* **6**, 625

— — 1937*b* *Act. Physicochim. U.S S.R.* **6**, 748.

Ku, Z. W. 1933 *Phys. Rev.* **44**, 383.

Lewis, B. and von Elbe, G. 1938 *Combustion flames and explosions in gases.*

Liefson, S. W. 1926 *Astrophys. J.* **63**, 73

Lord, R. C. and Wright, N. 1937 *J. Chem. Phys.* **5**, 642.

Mulliken, R. S. 1935 *J. Chem. Phys.* **3**, 720.

Price, W. C. and Simpson, D. M. 1938*a* *Proc Roy. Soc.* A, **169**, 501.

— — 1938*b* *Proc Roy Soc* A, **165**, 272.

Schmid, R. 1936 *Z. Phys.* **99**, 626.

Thompson, H. W. and Healey, N. 1936 *Proc. Roy. Soc.* A, **157**, 331.

Van Vleck, J. H. 1934 *J. Chem. Phys.* **2**, 20.

Voge, H H. 1936 *J. Chem. Phys* **4**, 581.

Weston, F. R. 1925 *Proc. Roy Soc.* A, **109**, 177, 523.

The high-frequency resistance of superconducting tin

By H. London
Wills Physical Laboratory, University of Bristol

(*Communicated by A. M. Tyndall, F.R.S.—Received* 11 *July* 1940)

The high-frequency resistance of tin in the superconducting state was measured at a wave-length of 20·5 cm. by a calorimetric method based on the principle of eddy-current heating. It was found that the resistance decreases *gradually* when the temperature falls below the transition point in contrast to the *sudden* drop in resistance peculiar to direct currents. An explanation of such a behaviour is given based on the assumption of the simultaneous presence of normal and superconducting electrons. Good agreement between theory and experiment was found.

Absolute measurements of the conductivity in the normal state at low temperatures with both high and low frequencies were carried out, and it was found that at the temperature of liquid helium the conductivity for high frequency is considerably lower than for low frequency. This behaviour is possibly due to the fact that the mean free path of the electrons becomes larger than the penetration depth due to skin effect under the conditions of high conductivity and high frequency.

Introduction

It has been pointed out in a previous paper (H. London 1934) that the possible simultaneous presence of superconducting and normal electrons would manifest itself by a resistance which is present for currents of very high frequency but absent for direct current. This resistance is due to the fact that the magnetic field associated with the electric current penetrates into a thin surface layer of the superconductor. If this magnetic field varies with time, an electric field will be induced and any normal conducting electrons which may be present will move in this field. Since they are subject to damping, they will produce Joule heat and give rise to an alternating current resistance.

Calculations on this basis show that if the number of superconducting electrons is comparable with the number of atoms, the alternating current resistance is exceedingly small except at very high frequencies.

After some preliminary measurements at $3 \cdot 7 \times 10^7$ and at $1 \cdot 5 \times 10^8$, the final measurements were carried out at $1 \cdot 46 \times 10^9$ cyc./sec., which corresponds to a wave-length of 20·5 cm.

APPARATUS

Most of the methods hitherto employed consisted in a purely electrical measurement of the resistance. As the resistance in question is very low, such methods are not very reliable. In particular, it is impossible to measure such low resistances *absolutely* at very high frequencies. The method used in the present experiments avoids the difficulties by measuring the Joule heat generated by a current induced in an ellipsoid of tin by an alternating electric field.

The calorimeter C consisted of a small pyrex glass vessel containing the tin specimen. It was thermally insulated by a vacuum jacket J which could be evacuated by the tube T. A capillary c connected the calorimeter to a mercury manometer M_1. Through this capillary the calorimeter could also be filled with liquid helium. If heat was generated in the tin, helium evaporated and the manometer rose. The rate of rise enables one to calculate the Joule heat, while the position of the manometer gives the vapour pressure of the helium in the calorimeter and hence its temperature. The manometer was read by a telescope with vertical traverse reading to $\frac{1}{100}$ mm. The position of the meniscus could be measured with an accuracy of 0·05 mm.

The vacuum jacket with the calorimeter was placed in a Dewar vessel D_1 which formed part of a small Linde liquefier.* It could be filled with liquid helium; this Dewar vessel was surrounded by another one, D_2, which was filled with liquid nitrogen. The lower parts of both these Dewar vessels were narrowed in order to fit into the oscillatory circuit which produced the alternating magnetic field. Heat radiation from outside was prevented from reaching the calorimeter by a screen S of black paper placed inside the liquid helium bath. This screen was painted with colloidal graphite, which, however, did not form a continuous layer, as otherwise it would have screened the high frequency. The glass of the Dewar vessels and the liquefier in the upper part of the inner Dewar gave additional protection against heat radiation.

The part of the capillary c which was at liquid helium temperature had an internal diameter of 0·3 mm. This diameter increased gradually to about 0·8 mm. in the room-temperature region. In this way the dead volume which is more effective at low temperatures was kept small, and at the same time the flow resistance which increases with temperature was not too high. It is particularly important to keep the capillary narrow inside the vacuum jacket, as the temperature of this part varies with the flow velocity of the evaporating helium. If the capillary is too wide, this causes irregularities of the temperature drift after the high frequency has been turned off, until temperature equilibrium is restored.

---

* A description of this liquefier, which departed from the all-metal technique hitherto employed, will be published elsewhere.

The high-frequency magnetic field was produced in an oscillatory circuit. It consisted of a brass tube of 6 cm. length, 3 cm. internal diameter, and 0·8 mm. wall, split at opposite sides by gaps parallel to the axis. The tube acted as a self-inductance, and the two gaps which were about 0·8 mm. wide formed condensers. The resonance wave-length of this arrangement was 20·5 cm. The ends of the split tube were surrounded by two short brass tubes of 3 cm. length, 4·7 cm. internal diameter and 1 mm. wall. These reduced the stray magnetic field of the oscillator and therefore diminished the radiation loss. One of the condenser gaps was connected to a tuned transmission line which led to a magnetron valve (G.E.C.-C.W. 10). The latter was sufficiently far distant to prevent the stray field of the magnet from affecting the superconductivity of the specimen. The line was bridged at a distance of one-quarter of a wave from the condenser gap, since otherwise the line would oscillate with a voltage node at the condenser gap at a frequency slightly different from the resonance frequency of the oscillatory circuit, as for all other frequencies the gap represents practically a short circuit.

The intensity of the oscillation was measured on an arbitrary scale by a vacuum junction and galvanometer, the former being connected to the other condenser gap by a quarter wave transmission line. The absolute value of the magnetic field for a given galvanometer deflexion was determined by calibration measurements at the temperature of liquid nitrogen, where the conductivity of tin is well known.

As an additional check the field was also determined by means of an ordinary mercury thermometer divided into tenths of a degree, which was inserted into the alternating magnetic field. The two methods agreed within about 10 %, which is within the limits of error.

The frequency was measured in the usual way by means of a Lecher wire system.

Outside the high-frequency circuit was placed a pair of Helmholtz coils by means of which an alternating field of 50 cyc./sec. could be produced. This was used for measuring the low-frequency conductivity of tin and for calibration purposes. The screening effect of the short brass tubes surrounding the ends of the oscillatory circuit was calculated and found to be negligible at 50 cyc./sec. Still further outside there was another pair of Helmholtz coils by which the earth's field could be compensated to less than 0·02 oersted.

The specimen was made of 99·996 % Hilger tin. It had the shape of an ellipsoid, its length was 3 cm. and its diameter 0·43 cm. It was melted *in vacuo* and solidified by raising the furnace slowly in order to avoid the formation of cavities. The sample obtained by this method consisted of a few large crystallites. A material of such structure was desirable as it transforms into the superconducting state in a smaller temperature interval than finely crystalline material (de Haas and Voogd 1931).

PROCEDURE

Since the Joule heat to be measured is very small (down to about 6×10^{-6} cal./min.), great care has to be taken in order to keep the temperature drift due to heat conduction small and regular. This can only be achieved if the vacuum inside the vacuum jacket is very good. It is impossible to obtain such a vacuum if the calorimeter is cooled in the usual way by filling the vacuum jacket with helium gas until the specimen has assumed the desired low temperature, and then pumping the gas away. A special technique suited to the present liquefier had therefore to be evolved. After the calorimeter had been cooled and filled, it was connected with the manometer M_1. The liquid helium still left in the capillary evaporated and the measurements could begin. These consisted in reading the manometer every half minute in order to measure the temperature drift due to heat radiation, switching on the high frequency for a period of 2–5 min. and then again taking the temperature drift until it had become linear again.

The gas taps of the manometer were always handled in such a way that gas which had been in contact with the mercury of the manometer could never enter the calorimeter, in order to avoid amalgamation of the specimen. Another precaution consisted in annealing the tin between the experiments at about 150° C for several hours in an atmosphere of helium or nitrogen, in order to counteract any possible transformation of the tin into its grey modification stable below 18° C. This might have occurred while the tin warmed up slowly after an experiment. The cooling always took place very rapidly.

EVALUATION OF THE MEASUREMENTS

In these measurements the helium does not evaporate at constant pressure. Therefore apart from the latent heat of evaporation the heat capacity of the calorimeter and the thermal expansion of the liquid and gaseous helium also enter into the relation between heat developed and the rate of rise of the manometer. The relation is somewhat complicated but can be expressed in terms of known data and is here omitted for the sake of brevity. The heat capacity of the glass container could be neglected. All calibrations of the calorimeter agreed to within 7 %.

The relation between the normal conductivity σ and the Joule heat q per sec. produced by a magnetic field of frequency ω and r.m.s. intensity H_0 is given by the following expression, which applies to small field penetration:

$$q = \frac{H_0^2}{8\pi}\sqrt{\left(\frac{\omega c^2}{2\pi\sigma}\right)}S, \tag{1}$$

where $$S = \frac{4\pi ab[\frac{1}{2}\epsilon\sqrt{(1-\epsilon^2)}+(\epsilon^2-\frac{1}{2})\arcsin\epsilon]}{\epsilon^3(1-A/4\pi)}$$

with the demagnetizing factor

$$A = 4\pi\frac{1-\epsilon^2}{\epsilon^2}\left(\frac{1}{2\epsilon}\ln\frac{1+\epsilon}{1-\epsilon}-1\right),$$

a and b are the semi-axes of the ellipsoid and ϵ its excentricity. S has the physical meaning of that part of the surface of an infinite cylinder in which the same heat q would be produced.

For the measurements at 50 cycles the approximation of small penetration is not applicable. Since the ellipsoid is rather long, a good approximation can be obtained by neglecting the derivatives of the field strength in the direction of the longer axis. One can then calculate the heat developed in any small disk perpendicular to the longer axis, as if the disk were part of an infinite cylinder having the radius of this particular disk. This problem can be solved rigorously and the total heat of the ellipsoid found by graphical integration. In the case of small penetration this method gives a result which agrees with (1) within 6·3 %, the discrepancy being mainly due to the fact that the demagnetizing effect is neglected in our approximation. Since this effect decreases with increasing penetration, the accuracy in cases where our approximation is used is much higher.

In the superconducting state the conductivity as defined by equation (1) has no physical significance. We shall represent the results in terms of the heat observed divided by the square of the applied magnetic field. This quantity is proportional to the high frequency resistance.

RESULTS

(a) *The transition curve*

In figure 2 the high-frequency resistance R divided by an average value of the resistance in the normal state just above the transition point R_n has been plotted against the absolute temperature (Leiden 1937 scale). At lower temperatures the points are plotted on a 10 times larger scale for R/R_n. Instead of the *sudden* decrease of resistance observed with direct current, we find a *gradual* decrease of R/R_n below the transition point. The value of the resistance observed at 2·3° K is only 0·6 % of the resistance in the normal state. From preliminary blank experiments made at a wavelength of 2 m., it is probable that the heat observed at 2·3° K is partly produced in the glass container. Unfortunately, the war made it impossible to make a blank run with the 20·5 cm. wave. Thus we have to take the value 0·006 at present as the upper limit of the resistance at that temperature.

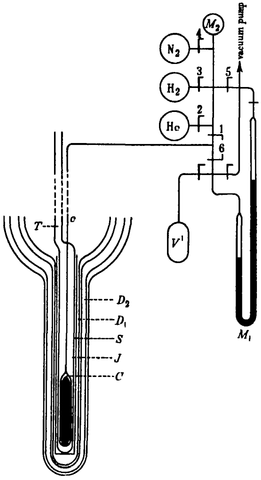


FIGURE 1

It is not probable that the heat due to the glass varies very much in the comparatively small temperature interval between 2·3 and 3·75° K. In order to minimize the uncertainty introduced by this effect, we have diminished all the results by $0{\cdot}003R_n$, that is, half the effect observed at 2·3° K. Then the error due to the glass will probably be nowhere greater than 0·5 % of the normal resistance. The measured points plotted in figure 2 are already corrected in this way. Apart from this constant error, there will be a proportional error of about 10 % for $T > 3{\cdot}5°$ K. Below this temperature the measurements are less accurate. The error is partly due to uncertainties in the drift correction and partly due to variations in the magnetron valve which did not settle down to stable conditions in the comparatively short heating periods The temperature interval covered during each heating period was always below 0·024°. The actual error in temperature is always less than 0·005° The peak value of the field was always less than 0·37 oersted, which would cause a decrease in transition temperature of only 0·003°.

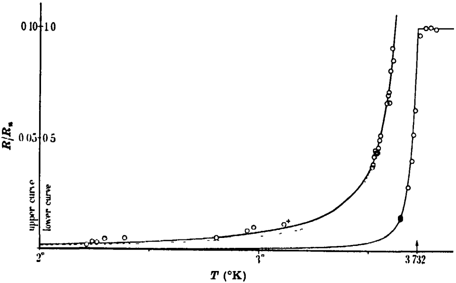


FIGURE 2

The point in figure 2 marked with a cross was obtained after the superconductivity had been temporarily destroyed by a strong magnetic field. This point lies less than 0·0015 above the adjacent point which was obtained before the field had been switched on. This is an indication that magnetic flux has been pushed out again nearly completely, since otherwise the remaining flux would have caused normal conducting inclusions which would give rise to an additional high-frequency resistance. As has been pointed out by Keeley and Mendelssohn (1936), complete expulsion of the lines of force is a criterion for high purity of the specimen.

The resistance begins to fall off about 0·01° above 3·723° K (vapour pressure 45·6 cm. Hg), which is the transition temperature of perfect single crystals of tin according to de Haas and Voogd (1931). This is in agreement with observations of these authors with specimens which like ours consisted of several large crystallites (*l.c.* figure 1; Sn 5–'30). With these they found an increase of the same order of magnitude in the temperature at which the first decrease in resistance occurred.

The high-frequency resistance observed in the present experiment is only about one-sixth of that observed by Burton and others (1933) with constantan wires coated with tin at a wave-length as long as 10 m. The difference is probably due to a penetration of the high frequency field into the constantan.

(*b*) *Conductivity in the normal state*

	σ in ohm^{-1} cm.$^{-1}$		ρ/ρ_0		
T °K	50 cyc./sec.	$1{\cdot}46 \times 10^9$ cyc./sec.	50 cyc./sec.	$1{\cdot}46 \times 10^9$ cyc./sec.	D C (Tuyn)
19	$1{\cdot}06 \times 10^7$	$0{\cdot}79 \times 10^7$	0·009	0·012	0·0102
3·8	$15{\cdot}0 \times 10^7$	$2{\cdot}06 \times 10^7$	0·00064	0·0046	0·0008

The table gives the absolute conductivity and the specific resistance in terms of the resistance at 0° C ($\rho_0 = 0{\cdot}105 \times 10^{-4}$ ohm cm.) at both low and high frequencies, and for comparison the direct current resistance found by Tuyn (1929). The accuracy of these measurements is about 15 % with the exception of the value at 3·8° K with 50 cyc./sec., the accuracy of which is only 25 %.* The table shows that there is good agreement with the D.C. values except for high frequency at helium temperatures, where the resistance is much higher than for low frequency. A small deviation in the same direction seems to be present also at 19° K.

DISCUSSION

(*a*) *The transition curve*

The simultaneous presence of superconducting and normal electrons referred to in the introduction can be represented by a modification of Ohm's law:

$$\dot{\mathbf{J}} = \sigma\dot{\mathbf{E}} + \frac{1}{\Lambda}\mathbf{E}, \tag{2}$$

---

* The accuracy is lower in this case because the Joule heat as a function of the conductivity has a maximum which for 50 cycles lies near the observed value of the conductivity.

where σ is the conductivity of the normal conducting electrons and Λ is a quantity taking account of the inertia of the superconducting electrons. Λ is connected with the thickness λ of the layer into which a stationary magnetic field penetrates by the expression

$$\Lambda = \frac{4\pi\lambda^2}{c^2}. \tag{3}$$

This formula was first introduced (H. London 1934) as a modification of the theory of Becker, Heller and Sauter (1933), who treated the superconducting electrons as free electrons which are accelerated by an electric field. Later, formula (2) could be incorporated in the phenomenological theory of F. and H. London (1935). Λ or λ define an effective number of free electrons per cm.^{3} by

$$N_S = \frac{m}{\Lambda e^2} = \frac{mc^2}{4\pi\lambda^2 e^2}, \tag{4}$$

where the electronic charge is measured in electrostatic units.

With the modified Ohm's law (2) the theory of quasistationary currents of the circular frequency ω can be developed in the usual way. Since the penetration depth is very small, the problem can be treated in one dimension. One obtains for the magnetic field strength at a depth, x, below the surface

$$H = H_0 e^{-zx}, \tag{5}$$

where $$z = \sqrt{\left\{\frac{1}{2}\left(\frac{1}{\lambda^2} + \sqrt{\left[\frac{1}{\lambda^4} + \frac{1}{\delta^4}\right]}\right)\right\}} + i\sqrt{\left\{\frac{1}{2}\left(-\frac{1}{\lambda^2} + \sqrt{\left[\frac{1}{\lambda^4} + \frac{1}{\delta^4}\right]}\right)\right\}}. \tag{6}$$

$$\delta = \sqrt{\frac{c^2}{4\pi\sigma\omega}} \tag{7}$$

is the length known from the ordinary theory of the skin effect where it is $\sqrt{\tfrac{1}{2}}$ times the penetration depth, while λ, as was mentioned above, is the penetration depth due to the superconducting electrons.

Analogous equations hold for the electric field and the current density.

From (5) we can deduce the following expression for the relative high-frequency resistance

$$\frac{R}{R_n} \equiv \frac{q}{q_n} = \frac{\lambda}{\delta_n}\sqrt{\frac{\sqrt{\{1+(\lambda/\delta)^4\}}-1}{1+(\lambda/\delta)^4}}. \tag{8}$$

In order to compare this expression with the measurements, we require the temperature dependence of λ and δ or, in other words, of the effective number N_S of superconducting electrons and of the conductivity σ due to the normal electrons. The quantity λ can be determined by experiments on thin films (Appleyard, Bristow, H. London and Misener 1939) and colloids (Shoenberg 1939, 1940), but so far only mercury has been investigated. Expressed by the effective number of electrons N_S as defined by (4), the results for the thin mercury films (Appleyard *et al.* 1939, figure 7), which are in good agreement with the colloid method, can be represented by the empirical expression

$$N_S/N_0 = 1-\tfrac{3}{4}(T/T_t)^2-\tfrac{1}{4}(T/T_t)^{16}, \tag{9}$$

with $N_0 = 0{\cdot}0644$ times the number of atoms per cm.^{3}. T_t is the transition temperature.

We now make the following assumptions:

(1) Equation (9) is also valid for tin, only the constants N_0 and T_t being different.

(2) The number of normal conducting electrons N_N is given by the equation

$$N_N = N_0 - N_S. \tag{10}$$

This implies the assumption that at absolute zero all electrons are superconducting, which is in agreement with the fact that the specific heat of superconducting tin contains no term proportional to the absolute temperature.

(3) The conductivity σ is given by

$$\sigma/\sigma_n = N_N/N_0, \tag{11}$$

where σ_n is the conductivity in the normal state. That is to say, it depends solely on the number of normal electrons and not explicitly on temperature or on the number of superconducting electrons. With respect to the temperature, this assumption seems to be justified in view of the fact that the conductivity of tin in the normal state is independent of temperature in the liquid helium region; the assumption that the superconducting electrons do not influence the normal ones is, of course, hypothetical.

Introducing (4), (7), (9), (10), (11) into (8) and using for σ_n the observed high-frequency value of $2{\cdot}06 \times 10^{-2}\, c^2$ (see table on p. 529), the only undetermined quantity is the number of superconducting electrons at absolute zero, N_0. The curve in figure 2 is drawn for $N_0/N_A = 0{\cdot}213$, where $N_A = 3{\cdot}90 \times 10^{22}$ is the number of atoms per c.c. of tin. It agrees with the observed points within the limits of error. The deviations close to the transition point are due to the fact mentioned above that the resistance begins to fall off about 0·01° above the transition point, while the calculated curve refers to an abrupt change of resistance at the transition point. The value of $N_0/N_A = 0{\cdot}213$ electrons per atom is of the order of magnitude which one would expect if the number of superconducting electrons at absolute zero is identical with the effective number of free electrons (see e.g. Fröhlich 1936, p. 190).

The good agreement between experiment and theory seems to confirm the hypothesis of the simultaneous presence of superconducting and normal electrons as expressed by the modified Ohm's law (3). But they should not be taken as a proof for our assumptions with respect to the temperature dependence of the number of superconducting electrons and to the normal conductivity expressed by the equations (9), (10), (11). For the main drop in resistance takes place within a region of 0·2° below the transition point where the normal conductivity has not yet decreased appreciably. Thus the drop in resistance is due rather to the appearance of the superconducting electrons than to the disappearance of the normal ones. At lower temperatures, on the other hand, the measurements are not sufficiently accurate. If one makes, e.g., the quite different assumptions that the normal conductivity does not vary at all, $\sigma = \sigma_0$, and that the number of superconducting electrons is proportional to $T_c - T$, one obtains a curve which between 3·5° K and the transition point is nearly coincident with the other curve; at lower temperatures it is shown by the dotted curve in figure 2.

At lower temperatures different assumptions about σ give of course very different results for N_S. Accordingly the value for N_0 given above has to be considered merely as an extrapolation made on the assumption that (9) holds for tin as well as for mercury.

On the other hand, quite definite statements can be made for the region near the transition point. First of all the experiments confirm the result found with thin films and colloids of mercury that the $N_S(T)$ curve meets the T-axis with a *finite slope*. As has been pointed out by Gorter and Casimir (1934), such a behaviour explains the discontinuity of the specific heat at the transition point. Furthermore, the value of this slope at the transition point can be determined by measurements within the main drop in resistance as it is not much affected by the particular assumptions about σ.

At still higher frequencies a further modification of Ohm's law is required, as with them the inertia of the normal electrons must not be neglected (H. London 1934, formula (5)). This does not affect the general feature of a gradual decrease in resistance to zero with falling temperature. In the infra-red, however, the conditions become different, when the wave-length is reached below which the absorption of energy is due mainly to optical transitions and not to the conductivity. Owing to the high value of normal conductivity at helium temperatures, this wave-length will be much larger than at room temperature, where it is of the order of magnitude of 20μ. This is in agreement with the observations of Daunt, Keeley and Mendelssohn (1937) with lead and tin, that there is no change in the absorption of infra-red light of a wave-length of about 10–20μ, when these metals pass into the superconducting state.

(*b*) *High-frequency resistance in the normal state*

The anomalous normal resistance at high frequencies can probably be explained by the fact that the mean free path of the electrons is considerably larger than the penetration depth An increase of resistance due to this effect has been observed by Lovell (1936) and Appleyard and Lovell (1937) in experiments on thin alkali metal films.

This work was begun at the Clarendon Laboratory, Oxford and continued at the Wills Physical Laboratory, Bristol. I should like to express my thanks to Professor F. Simon for many valuable suggestions and to Professor A. M Tyndall for his support and interest. I am very grateful to Dr L. C Jackson for his advice and for supplying the liquid hydrogen, and to him and Dr A. D. Misener for their help during the measurements.

REFERENCES

Appleyard, E. T. S., Bristow, J. R., London, H. and Misener, A. D. 1939 *Proc. Roy. Soc.* A, **172**, 540.

Appleyard, E. T. S. and Lovell, A. C. B. 1937 *Proc. Roy. Soc* A, **158**, 718

Becker, R., Heller, G. and Sauter, F. 1933 *Z. Phys.* **85**, 772.

Burton, E F., Wilhelm, J. O , Pitt, A. and Young, A. C. 1933 *Canad. J. Res* **9**, 630.

Daunt, J G., Keeley, T. C. and Mendelssohn, K. 1937 *Phil. Mag.* **23**, 264

Frohlich, H. 1936 *Elektronentheorie der Metalle*, p. 190. Berlin Springer.

Gorter, C. S. and Casimir, H. 1934 *Phys. Z.* **35**, 963.

de Haas, W. J. and Voogd, J. 1931 *Commun. Phys. Lab Univ. Leiden*, no. 214C.

Keeley, T. C. and Mendelssohn, K 1936 *Proc Roy. Soc.* A, **154**, 378.

London, F. and H. 1935 *Physica*, **2**, 341.

London, H. 1934 *Nature, Lond* , **133**, 497.

Lovell, A. C. B. 1936 *Proc. Roy. Soc.* A, **157**, 311.

Shoenberg, D. 1939 *Nature, Lond* , **143**, 434.

— 1940 *Proc. Roy. Soc.* A, **175**, 49.

Tuyn, W. 1929 *Commun Phys. Lab. Univ. Leiden*, no. 196B.

On the structure of multilayers. Part I

By G. Knott, J. H. Schulman and A. F. Wells

Department of Colloid Science and Crystallographic Laboratory, Cambridge

(*Communicated by E. K. Rideal, F.R.S.—Received* 11 *July* 1940)

[Plate 15]

X-ray and microscopic examinations of multilayers of certain long-chain esters deposited on transparent bases have been made. An X-ray photograph of a stationary multilayer of ethyl stearate is apparently identical with that obtained by rotating a single crystal of β-ethyl stearate about the long axis of the cell. The spots previously observed on oscillation photographs of multilayers of ethyl stearate are in reality parts of circles, successive arcs of which appear on increasing the angle of incidence of the X-rays on the multilayer. These facts reveal that the multilayer consists of microcrystals all oriented with one axis in common, this axis being the normal to the surface of the multilayer. The single crystals appear to have grown without interruption throughout the entire thickness of the multilayer, which in these investigations was about 1000 mol. thick.

Whatever be the structure of multilayers of long-chain esters at the moment of deposition, they consist after a certain interval of microscopic crystals. This was established by Stenhagen (1938) for certain long-chain esters. His observations were made on multilayers deposited on metal bases and therefore may only refer to surface structure. The correspondence between the long spacings of the multilayers, irrespective of the type of deposition (X, Y, Z), and those of some of the known crystalline forms of the esters indicates that these crystals may be similar to those of certain polymorphic forms of the substances. Owing to the fact that the side-spacings of neither the multilayers nor the corresponding esters have yet been measured we have no knowledge of the arrangement of the molecules in the multilayers. If we assume that the crystals in the multilayer are those of one of the forms of the crystalline ester then there are two problems to be solved. The first is to discover the arrangement of these crystals in the multilayer, and the second is to determine the actual molecular structure of the crystals themselves. The latter is a purely crystallographic problem which is conveniently dealt with separately. The former presents two questions: (1) whether the surface structure persists through the multilayer, i.e. whether a given portion is a single crystal from the base to the surface of the multilayer, and (2) how the single crystals are oriented in relation to one another.

By depositing the multilayers on transparent bases it has been possible to examine them microscopically in transmitted light. Examination between crossed Nicols showed that the individual crystals extinguished perfectly. This indicates that the molecules are arranged in some regular way from the base of the multilayer to the surface, i.e. that the grain boundaries seen on the surface are those of single crystals which have grown without interruption through the entire thickness of the multilayer. A recent paper suggests an answer to the second problem. Bernstein (1940) has taken an oscillation photograph of an ethyl stearate multilayer built up on a metal base. This shows a single side-spacing reflexion, from which the author deduces that in the case of this multilayer all the crystals have the same orientation. This deduction from the photograph is not justifiable, since it implies that the multilayer is a normal mosaic crystal. The only conclusion which can be drawn from the presence of this discrete reflexion is that the multilayer does not consist of crystals arranged entirely at random.

The experiments to be described show that the side-spacing reflexion observed by Bernstein is in reality not a spot, as would be the case from a single crystal, but part of a circle. The rest of the circle and also parts of others appeared on increasing the angle of incidence of the X-rays on the multilayer until finally a transmission photograph was obtained. The taking of such photographs has been made possible by depositing the multilayer on a celluloid instead of a metal base. These photographs, which are reproduced, show that the small crystals composing the multilayer are oriented with one axis in common, and this interpretation has been fully confirmed. The following account deals with observations made on multilayers of ethyl and methyl stearates and octadecyl acetate. The X-ray examination of multilayers of the last compound led to the same conclusions as for ethyl stearate The work on ethyl stearate is described in greater detail because by far the best photographs were obtained from ethyl stearate multilayers and also because a crystallographic examination of single crystals of this compound is in progress, and it is hoped that this will make possible a complete interpretation of the multilayer photographs.

Before proceeding to give the detailed results it should be pointed out that all the conclusions, like those of other X-ray workers in this field, refer to the structure of the multilayers some time after deposition, since the time required to take an X-ray photograph is of the order of several hours. We have observed that the freshly deposited multilayers show no microcrystalline structure, this appearing after ½–1 hr. It is hoped to investigate the detailed changes which may take place in the multilayer immediately after deposition. There may, for example, be loss of water followed by recrystallization, for it must be remembered that we are working at temperatures very near the melting points of these compounds (e.g. ethyl stearate, m.p. 33·5°), and molecular rearrangement may take place easily at room temperature.

Preparation of the multilayers

The monolayers of the esters were deposited from distilled water on to thin transparent sheets of various materials, celluloid, cellulose acetate, perspex, polystyrene, formvar, vinyl acetates and other synthetic resins, by the method described by Blodgett (1935) and also by means of a mechanical dipper. Celluloid sheets were found to give the most satisfactory results, for the X-ray diffraction of this base interfered least with the interpretation of the photographs. The small celluloid strips, about 1 cm. square, were held on both sides of a chromium-plated holder by means of pure vaseline. The multilayer was thus deposited on only one side of the celluloid strip which could be easily removed from the metal base after deposition of the monolayers and placed in supports for X-ray and microscopic examination. This enabled photographs to be taken for any angle of incidence of the X-ray or light beam. X-ray photographs of multilayers deposited mechanically and by hand appeared to be identical. Films of ethyl stearate and of octadecyl acetate were deposited under a piston pressure of 31 dynes/cm. (oleic acid) and those of methyl stearate under a piston pressure of 16 dynes/cm. (castor oil); all gave *Y*-deposition. All traces of dust and small crystallites from possible excess material on the surface or from collapsed films were carefully avoided. The compounds used were those described by Alexander and Schulman (1937).

Microscopic examination

Ethyl stearate. Examination of the film with high magnification between crossed Nicols showed the film to consist of very small single crystals, each with a definite extinction direction. Detailed examination of a small portion of the multilayer showed that these extinctions are grouped about two mean directions inclined at approximately 45° to one another. This is seen from a pair of photomicrographs taken between crossed Nicols for two positions of the Nicols relative to the multilayer. The two positions of the optic axis of the analyser with respect to the dipping direction are shown in figure 1, and in plate 15 it is seen that the photographs c and e are approximately complementary, i.e. the portions which are black in one are white in the other. If the extinction directions of *all* the crystals composing the multilayer lay along either a or b this would be equivalent to simple twinning. However, measurement of the extinction directions of the individual crystals shows that they are distributed over a wide range of angle on each side of the mean directions a and b. For this reason, presumably, the X-ray photographs do not show this twinning effect, but it is sufficiently striking in the photomicrographs to be worth recording.

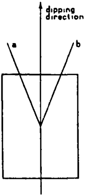


FIGURE 1. The mean extinction directions (a, b) of the microcrystals in an ethyl stearate multilayer.

Methyl stearate. Photomicrographs of a methyl stearate multilayer show an entirely different structure. Between crossed Nicols much of the field was extinguished at the same time and the intermediate portions were brought into the extinction position by a rotation of the Nicols relative to the multilayer of only 18° (plate 15 a, b). In contrast to the well-defined single crystals in the ethyl stearate multilayer methyl stearate shows larger areas extinguishing uniformly, and on rotation of the Nicols the dark areas move across the field, suggesting a gradual change in orientation of the crystallites in neighbouring parts of the multilayer. These effects and the striations parallel to the dipping direction suggest that the crystallization of the multilayer has proceeded differently from that of ethyl stearate, and the X-ray photographs were inferior to those of ethyl stearate multilayers. For this reason the latter have been studied in greater detail. In plate 15 an enlargement of c is shown (d) to emphasize the difference in microcrystalline structure between the methyl and ethyl stearate multilayers.

Octadecyl acetate. Between crossed Nicols multilayers of this compound were seen to be composed of large irregular patches most of which extinguished perfectly for a particular setting of the Nicols as in the case of ethyl stearate. Other patches showed a dark brush which travelled across the patch as the Nicols were rotated with respect to the multilayer. One or two specimens showed very small needle-shaped crystals oriented at random, but these were probably due to partial melting and recrystallization on the surface.

X-RAY EXAMINATION

Ethyl stearate. A series of X-ray photographs have been taken of ethyl stearate multilayers (approximately 1000 mol. thick) using a Metallix tube with copper anti-cathode running at 25 mA, the beam being monochromatized by reflexion from a crystal of pentaerythritol. The multilayer was first oscillated through 15° about an axis in its own plane, the X-rays being parallel to the surface at the beginning of the oscillation. The dipping direction was chosen as a suitable axis. Plate 15 *f* shows the photograph (0–15°) obtained with an exposure time of 6 hr., the flat plate being at a distance of 4 cm. from the axis of oscillation. Apart from the first seven main orders there is a layer line of six spots at $\zeta = 0{\cdot}34$ and one other reflexion at $\xi = 0{\cdot}09$, $\zeta = 0{\cdot}41$. The spacing of the main orders (25·2 A), which is approximately that of single crystals of β-ethyl stearate (Malkin 1931), shows that we are dealing with the same type of multilayer as examined by previous workers. The multilayer was then tilted in its own plane and this photograph retaken. It was found that the same photograph is obtained by oscillating about any axis in the plane of the multilayer. A set of 15° oscillation photographs was then taken starting with the X-rays parallel to the surface (0–15°) and ending with the X-rays passing normally through the multilayer (75–90°). These photographs show that all the spots except the main order reflexions are portions of circles, consecutive arcs of which appear as the angle between the X-ray beam and the normal to the surface decreases. The interpretation of the photographs is most simply given in terms of the reciprocal lattice in which a set of parallel planes of spacing d is represented by a point at a distance λ/d from the origin along the normal to the planes. The reciprocal lattice of the multilayer may be reconstructed from the photographs. Figure 2 (*a*) shows the projection of this lattice on a plane normal to that of the multilayer, and figure 2 (*b*) the projection on the plane of the multilayer. The reciprocal lattice of a single crystal consists of points forming a three-dimensional array, that of an aggregate of single crystals oriented at random consists of spheres around the origin. The X-ray photograph in the latter case is the powder photograph which, if taken on a flat plate, is a series of concentric circles. In this multilayer we have an intermediate case. Instead of the crystals being oriented entirely at random they all have one axis in common owing to the fact that the angle of tilt of the molecules to the surface is constant. The reciprocal lattice of the multilayer is therefore the result of rotating the lattice of the single crystal about a line. The main orders lie on this line (the normal to the surface of the multilayer) and therefore remain points, but all other points in the lattice become circles. Accordingly, if the multilayer is oscillated through a limited angle, arcs of circles appear on the photographs. Moreover, even with monochromatic radiation and with the specimen stationary spots are obtained on a photograph since the sphere of reflexion cuts a number of circles in the reciprocal lattice. As a check on this interpretation the portions of each circle which should reflect on a given photograph were deduced graphically from the reciprocal lattice. These corresponded precisely to the photographs of plate 15 *f*.

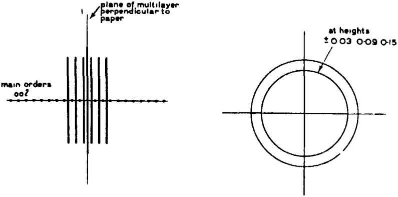


FIGURE 2. (*a*) Portion of the reciprocal lattice of the ethyl stearate multilayer projected on a plane perpendicular to that of the multilayer. The heavy lines are the projections of circles and give rise to the reflexions, other than the main orders, in the photographs of plate 15 *f*. (*b*) Projection of the same lattice on the plane of the multilayer.

The spacings of the strongest side-spacings are 3·85 and 4·15 A. The weaker intermediate reflexions on the layer line correspond to spacings of 4·38 and 4·51 A. The main orders (00*l*) show interesting intensity variations. The intensities fall off suddenly after the first four orders, which in relative intensities have the following sequence· 002 > 003 > 001 > 004. The next four orders are weak, and 009 and 00.10 were not observed with an exposure time of 5–6 hr. The reflexion 00.11 is strong and 00.12 very weak. The appearance of the strong eleventh order is due to the fact that alternate carbon atoms of the chains are in phase for this reflexion. Assuming the C–C distance to be 1·54 A and the tetrahedral angle between the carbon bonds we have

$$\sin\theta = \frac{c(=25{\cdot}2)}{11\times 2(1{\cdot}54\cos 35^\circ)},$$

whence the angle of tilt (θ) of the chains to the surface may be calculated as 65°. The presence of 00.12 suggests that this angle may be rather less than 65°. In addition to the main order reflexions corresponding to $c = 25{\cdot}2$ A,† two other reflexions have been observed in the same zone. We have not yet found a satisfactory explanation for the presence of these two spots. One, which is very weak, lies midway between 002 and 003 and therefore corresponds to a spacing of 10·08 A. It will be noticed that this point in the reciprocal lattice lies in the plane of the circle corresponding to the strongest side-spacing. Normally the presence of such a 'halving' would be taken to indicate that the strong reflexions are the even orders from planes of spacing 50·4 A instead of 25·2 A. The doubling of this dimension would also account for the fact that the projections of the circles cross the c^* axis midway between the main orders, but a non-orthogonal reciprocal lattice (with suitable interaxial angles) rotated about c^* would equally well explain the positions of these circles. Moreover, the doubling of the c spacing does not explain the presence of the second anomalous reflexion which lies between, though not midway between, the sixth and seventh orders and is appreciably stronger than either of these. Its spacing is 4·15 A. It must be recorded that there is no trace of this reflexion on one photograph of a freshly prepared multilayer

The photographs so far described were taken with exposures of the order of 150 mA/hr. Subsequently the exposure time was increased and very beautiful photographs have been obtained after 750–1000 mA/hr. The presence of circles in the reciprocal lattice having been established, it is not necessary to oscillate the specimen, and one photograph summarizes all the available data concerning the side-spacings A tracing of such a photograph is shown in figure 3. (A tracing is shown instead of the actual photograph since many of the reflexions are weak and would be lost on reproduction.) This photograph was taken in a cylindrical camera (radius 3 cm.) with the multilayer horizontal and the X-rays parallel to the surface. The innermost row-lines therefore correspond to the layer-lines on the first photograph of plate 15 *f*.

The final proof of the structure of the ethyl stearate multilayer comes from a rotation photograph of a single crystal of the ester about the long axis of the cell. The photograph appears to be identical with the original of figure 3. The detailed interpretation of this photograph will be dealt with in Part II. It is clearly impracticable to attempt to reconstruct the reciprocal lattice of the single crystal from that of the multilayer since all that we know about the position of a point in the former is that it lies somewhere on the circumference of the circle in the latter reciprocal lattice. Since we are dealing with single crystals of β-ethyl stearate it is far simpler to study single crystals of this compound and then to reconstruct the reciprocal lattice of the multilayer from that of β-ethyl stearate.

---

† The value 25·7 A for this spacing recorded by Stenhagen (1938) appears to be too high. Malkin (1931) gives 25·5 A for the crystalline ester and Bernstein (1940) 25·2 A for the multilayer.

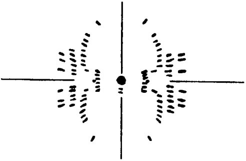


FIGURE 3. Tracing of an X-ray photograph of an ethyl stearate multilayer.

Octadecyl acetate. Stenhagen (1938) has shown that there are two forms of built-up multilayers of octadecyl acetate—a 'single-layer' form of 28·7 A spacing and a 'double-layer' form of 39·0 A spacing. In the latter form the long axes of the molecules are inclined at 44° to the surface of the multilayer. X-ray photographs of this form have now been taken about axes in the plane of the multilayer and about the normal to the surface. Orders of the long spacing up to the 22nd were observed, this reflexion being strong owing to the carbon atoms scattering almost in phase. The spacing of this plane, 1·77 A, confirms the value of 44° for the tilt of the molecules. In addition to the main orders other reflexions were observed, apparently lying on layer-lines. As in the case of ethyl stearate subsequent investigation showed that these reflexions are parts of continuous rings, so that the multilayer of octadecyl acetate clearly has a similar structure to that of ethyl stearate. A photograph of one multilayer showed also faint but quite definite layer-lines corresponding to a side-spacing of 4·2 A. On the first layer-line some ten reflexions could be seen, and on the second three weak reflexions. The separation of these spots along the layer-line indicates that the true c spacing of the single crystals is 78·0 A. The main orders are therefore the even orders. The appearance of these genuine layer-lines was apparently due to the presence of one single crystal in the multilayer much larger than those normally observed.

Our thanks are due to Professor E. K. Rideal, F.R.S., for much help and stimulating discussion, to the Medical Research Council for a personal grant (J.H.S.), and to the Department of Scientific and Industrial Research for a Senior Research Award (A.F.W.).

REFERENCES

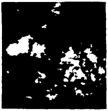
Alexander, A. E. and Schulman, J. H. 1937 *Proc. Roy. Soc.* **A, 161, 115.**
Bernstein, S. 1940 *J. Amer. Chem. Soc.* **62, 324.**
Blodgett, K. B. 1935 *J. Amer. Chem. Soc.* **57,** 1007.
Malkin, T. 1931 *J. Chem. Soc.* p. **2796.**
Stenhagen, E. 1938 *Trans. Faraday Soc.* **34, 1328.**

DESCRIPTION OF PLATE 15

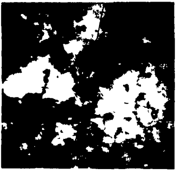
a, *b*, Photomicrographs of methyl stearate multilayer; *c*, *d*, *e*, photomicrographs of ethyl stearate multilayer (see p. 537); *f*, a series of 15° X-ray oscillation photographs of an ethyl stearate multilayer starting with the X-rays parallel to the surface (0°) and finishing with the X-rays normal to the surface (90°).

a

b

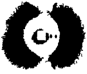


c

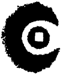


d

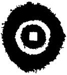
e



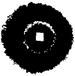
0–15°



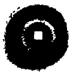
15°–30°



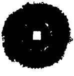
30°–45°



45°–60°



60°–75°



75°–90°

(*Facing p.* 542)

INDEX TO VOLUME 176 (A)

Acetones, substituted, based catalysed prototropy (Bell and Lidwell), 88.

Acetophenone, thermal decomposition (Hinshelwood and Smith), 468

Active nitrogen, studies, I, II (Rayleigh), 1, 16.

Atmospherics at night, wave form (Elder, Hodges, Phillips, Schonland and van Wyk), 180.

Austen, A. E. W. and Whitehead, S. The electric strength of some solid dielectrics, 33.

Autelectronic emission, distribution from single crystal metal points (Benjamin and Jenkins), 262.

Bakerian Lecture. Stereochemical types and valency groups (Sidgwick and Powell), 153.

Bell, R. P. and Lidwell, O. M. The base catalysed prototropy of substituted acetones, 88.

Bell, R. P. and Lidwell, O. M. Potential energy curves in proton transfer reactions, 114.

Benjamin, M. and Jenkins, R. O. The distribution of autelectronic emission from single crystal metal points, I, 262.

Binary and ternary systems, two-phase equilibrium, II, III (Guter, Newitt and Ruhemann), 140.

Binary and ternary systems, two-phase equilibrium, IV (Burgoyne), 280.

Burgoyne, J. H. Two-phase equilibrium in binary and ternary systems, IV, 280.

Burrows, M. G. T. and Stockmayer, W. H. The poisoning of a palladium catalyst by carbon monoxide, 474.

Carbon monoxide, flame spectrum, 505.

Chao, S. H. and Taylor, W. H. Isomorphous replacement and superlattice structures in the plagioclase felspars, 76.

Clews, C. J. B. and Robinson, H. R. X-ray electrons expelled from metals by silver $K\alpha_1$ radiations, 28.

Combustion of hydrocarbons, surface as a limiting factor (Norrish and Reagh), 429.

Conn, G. K. T., Lee, E., Sutherland, G. B. B M. and Wu, Ch.-K. Investigations on the vibration spectra of certain condensed gases at the temperature of liquid nitrogen, I, 484.

Convective motion in a fluid heated from below (Pellew and Southwell), 312.

Dielectrics, solid, electric strength (Austen and Whitehead), 33.

Elder, J. S., Hodges, D. B., Phillips, W. E., Schonland, B. F. J. and van Wyk, J. W. The wave form of atmospherics at night, 180.

Fuchs, K. Operator calculus in the electron theory of metals, 214.

Gaydon, A. G. The flame spectrum of carbon monoxide, 505.
Green, A. E. General bi-harmonic analysis for a plate containing circular holes, 121.
Guter, M., Newitt, D. M. and Ruhemann, M. Two-phase equilibrium in binary and ternary systems, II, III, 140.

Heitler, W. and Ma, S. T. Inner excited states of the proton and neutron, 368.
Hinshelwood, C. N and Smith, R. E. The kinetics of the thermal decomposition of acetophenone, 468.
Hodges, D. B. *See* Elder and others.
Jenkins, R. O. *See* Benjamin and Jenkins.

Kendrew, J. C and Moelwyn-Hughes, E. A. The kinetics of mutarotation in solution, 352.
Kinetics of mutarotation in solution (Kendrew and Moelwyn-Hughes), 352.
Knott, G., Schulman, J. H. and Wells, A. F. On the structure of multilayers, I, 534.

Lee, E., Sutherland, G. B. B. M. and Wu, Ch.-K. Investigations on the vibration spectra of certain condensed gases at the temperature of liquid nitrogen, II, 493.
Lee, E. *See also* Conn and others.
Lidwell, O. M. *See* Bell and Lidwell
London, H. The high-frequency resistance of superconducting tin, 522.

Ma, S. T. *See* Heitler and Ma.
Melting and crystal structure (Oldham and Ubbelohde), 50.
Micelle formation in colloidal electrolytes, influence of the solvent (A. F. H Ward), 412
Moelwyn-Hughes, E A *See* Kendrew and Moelwyn-Hughes
Multilayers, structure (Knott, Schulman and Wells), 534

Newitt, D. M. *See* Guter, Newitt and Ruhemann.
Norrish, R. G. W. and Reagh, J. D. The surface as a limiting factor in the slow combustion of hydrocarbons, 429.
Norrish, R G. W and Smith, W. MacF. The quenching of the resonance radiation of sodium, 295.

O'Bryan, H. M. and Skinner, H. W. B. The soft X-ray spectroscopy of solids, II, 229.
Oldham, J. W. H. and Ubbelohde, A. R. Melting and crystal structure, 50.
Operator calculus in the electron theory of metals (Fuchs), 214.

Palladium catalyst, poisoning by carbon monoxide (Burrows and Stockmayer), 474.
Paramagnetic alums at low temperatures (Sauer and Temperley), 203.
Pellew, A. and Southwell, R. V. On maintained convective motion in a fluid heated from below, 312.
Phillips, W. E. *See* Elder and others.

Plagioclase felspars, isomorphous replacement and superlattice structures (Chao and Taylor), 76.
Plate containing circular holes, general bi-harmonic analysis (Green), 121.
Powell, H. M *See* Sidgwick and Powell.
Proton transfer reactions, potential energy curves (Bell and Lidwell), 114
Proton and neutron, inner excited states (Heitler and Ma), 368.
Proudman, J On the turbulence of a tidal current, 449.

Quenching of the resonance radiation of sodium (Norrish and Smith), 295.

Rayleigh, Lord. New studies on active nitrogen, 1, 16.
Reagh, J. D. *See* Norrish and Reagh
Robinson, H. R. *See* Clews and Robinson.
Ruhemann, M. *See* Guter, Newitt and Ruhemann.

Sauer, J. A. and Temperley, H. N. V. A theoretical study of a possible model of paramagnetic alums at low temperatures, 203.
Schonland, B. F. J. *See* Elder and others.
Schulman, J. H. *See* Knott, Schulman and Wells
Sidgwick, N. V. and Powell, H. M. Bakerian Lecture. Stereochemical types and valency groups, 153.
Skinner, H. W. B *See* O'Bryan and Skinner.
Smith, R. E. *See* Hinshelwood and Smith.
Smith, S. L. and Wood, W. A. X-ray structure and elastic strains in copper, 398.
Smith, W. MacF. *See* Norrish and Smith.
Soft X-ray spectroscopy of solids, II (O'Bryan and Skinner), 229.
Southwell, R. V. *See* Pellew and Southwell.
Srivastava, B. N. Thermal ionization of strontium, 343.
Stereochemical types and valency groups (Bakerian Lecture) (Sidgwick and Powell), 153.
Stockmayer, W. H. *See* Burrows and Stockmayer.
Superconducting tin, high-frequency resistance (London), 522.
Sutherland, G. B. B. M. *See* Conn and others *and* Lee and others.

Taylor, W. H. *See* Chao and Taylor.
Temperley, H. N. V. *See* Sauer and Temperley.
Thermal ionization of strontium (Srivastava), 343.
Turbulence of a tidal current (Proudman), 449.

Ubbelohde, A. R. *See* Oldham and Ubbelohde.

Vibration spectra of condensed gases (Conn, Lee, Sutherland and Wu), 484.
Vibration spectra of hydrochloric and deuterochloric acid at liquid nitrogen temperatures (Lee, Sutherland and Wu), 493.

Ward, A. F. H. The influence of the solvent on the formation of micelles in colloidal electrolytes, I, 412.
Wells, A. F. *See* Knott, Schulman and Wells.

Whitehead, S. *See* Austen and Whitehead.
Wood, W. A. *See* Smith and Wood.
Wu, Ch.-K. *See* Conn and others *and* Lee and others.
van Wyk, J. W. *See* Elder and others.

X-ray electrons expelled from metals by silver Kα_1 radiations (Clews and Robinson), 28.
X-ray structure and elastic strains in copper (Smith and Wood), 398.

END OF THE ONE HUNDRED AND SEVENTY-SIXTH VOLUME (SERIES A)

ERRATUM

On plate 4 for *Facing p.* 245 read *Facing p* 264

ABSTRACTS

OF PAPERS COMMUNICATED TO THE ROYAL SOCIETY OF LONDON

In accordance with a resolution of Council, summaries or abstracts of papers are to be published as soon as practicable. The publication of such abstracts in no way indicates that the papers have been accepted for publication in any fuller form. These abstracts are issued for convenience with the "Proceedings of the Royal Society of London" but do not form a part of the "Proceedings".

---

9 August 1940

A spectroscopic investigation of hydrocarbon flames. Parts II and III.

By W. M Vaidya (*Communicated by M. N. Saha, F.R.S.—Received 15 May* 1940.)

The flames of hydrocarbons burning in atomic oxygen have been re-examined in order to find out whether the ethylene-flame bands which occur extensively in the flames of carbon compounds burning in air also occur in such atomic flames. The results show that ethylene, acetylene and benzene which gave the ethylene-flame bands burning in air also yield them when burnt in atomic oxygen. The intensities are, however, different Methane does not react with atomic oxygen Methyl alcohol gives CH, HO and the cool-flame spectrum of ether, while formaldehyde shows only HO bands.

The ethylene-flame bands are absent from the flame of benzene burning in atomic hydrogen, which means that oxygen is essential for their production. An emitter like CH_2 is, therefore, not possible.

The results obtained in the study of the spectra of hydrocarbon flames and the flames of intermediate products, burning in air and atomic oxygen, are discussed. It is suggested that, in hydrocarbon flames, the C_2 bands may be produced through the decomposition of a higher complex of carbon and oxygen such as C_3O_2 and the ethylene-flame bands from the decomposition of an oxy-compound formed by the direct incorporation of an oxygen molecule in a hydrocarbon molecule.

The nature of the emitter of the ethylene-flame bands is discussed.

The general problem of the mechanism of hydrocarbon combustion is discussed in the light of the spectroscopic data. It is concluded that the chain mechanism in which oxygen atoms had been assumed as chain carriers does not seem very likely; the evidence appears to be more in favour of the peroxidation theory.

The surface as a limiting factor in the combustion of hydrocarbons.

By R. G. W. Norrish, F.R.S. and J. D. Reagh. (*Received* 17 *May* 1940.)

With special reaction vessels, varying widely in diameter while maintaining approximately constant volume, and employing several refinements in technique, a study has been made of the effect of surface on the slow oxidation of several hydrocarbons, both saturated and unsaturated. All reactions were of the degenerate branching type, and were found to be principally homogeneous in character. When the diameter of the reaction vessel was sufficiently reduced, the reaction rate was observed to drop abruptly toward zero, while the corresponding induction period increased toward infinity. Data so obtained have demonstrated that in narrow vessels surface deactivation can predominate over other processes of deactivation, while in wider vessels surface and volume deactivation occur to a comparable extent over a considerable range of pressure. When the vessel diameter is decreased to a critical value the surface deactivation almost alone can suppress the factors leading to chain branching, and from the reactions investigated the existence of such a critical diameter appears to be a general property of hydrocarbon oxidations in conformity with the theory of degenerate branching.

'Traces' of frictional contact and experiments on adhesion.

By R. Schnurmann. (*Communicated by Sir Harold Hartley, F.R.S.—Received* 4 *June* 1940.)

'Traces' due to frictional contact in the atmosphere can be made visible on glass plates through physical development by the condensation *in vacuo* of a metallic vapour. Above a certain load traces were obtained with silver, copper, steel, lead and cadmium.

The essential difference between traces of sliding and traces of rolling was demonstrated through microscopic observation Traces of sliding are resolved into wide bands of predominantly parallel lines, whereas traces of rolling consist of numerous isolated specks. The minimum load required for the production of a trace of rolling is larger than for producing a trace of sliding. A more sensitive indicator of traces of rolling was arrived at by predeveloping glass plates

The production of a trace on a glass plate is accompanied by the corresponding generation of a trace on the metal surface. This latter trace can be observed microscopically on the originally polished surface of a metal slider. Intentional lubrication, such as flooding the glass plate with distilled water or application of stearic acid, accentuates the traces.

Metals cast on to glass plates exhibit strong adhesion, unless they are appreciably superheated before casting. The contact area appears milky and full of blowholes, whilst the islands between the blowholes show an etched structure.

Evidence is produced that water vapour acts as an etching agent for the adhering metal surfaces. The adsorption layer trapped at the areas of 'contact' between etched metal and glass is considered to be sufficiently thin to be adhesive. Oxygen etching leads to a different result, since in this case the metal cracks off the glass plate on solidification leaving a trace behind.

The theory that local melting can occur at surface asperities on frictional contact is critically examined. Traces due to sliding, rolling, or strong adhesion were found under conditions which illustrate the importance of very thin adsorption and oxide layers, which the author considers cannot be reconciled with the existence of "welded metallic bridges".

Experiments on the elementary mechanism of solid friction. By R. Schnurmann. (*Communicated by Sir Harold Hartley, F.R.S.—Received 4 June* 1940.)

If deformation of surface irregularities were the predominant mechanism of solid friction, the coefficient of friction should have a measurable temperature coefficient. Apparatus is described, embodying the principle of the inclined plane, on which experiments in high vacuum were carried out over a wide range of temperature to determine whether such a coefficient could be detected.

Rust and adsorbed layers were found greatly to influence the measurements and a technique of cleaning the friction elements by prolonged baking out was developed. Thereafter, experiments on cadmium in the range between + 150 and − 100° C clearly foreshadow the existence of a temperature coefficient of static friction, computed at $-7{\cdot}5 \times 10^{-4}$ per $+1^\circ$ C.

Two-phase equilibrium in binary and ternary systems. IV. The thermodynamic properties of propane. By J. H. Burgoyne. (*Communicated by A. C. G. Egerton, Sec. R.S.—Received* 4 *June* 1940.)

Existing physical and thermal data relative to propane have been summarized and correlated, and some new experimental determinations of pressure-volume-temperature relationships for the liquid at low temperatures have been carried out to make good deficiencies in the literature. On the basis of the information thus obtained the entropy and enthalpy of propane have been calculated for conditions of temperature between − 80 and 200° C, and at pressures of from 0·1 to 200 atm. The results are tabulated and also presented graphically on a temperature base.

The estimation of small quantities of carbon dioxide in air by the absorption of infra-red radiation. By H. Dingle and A. W. Pryce. (*Communicated by V. H. Blackman, F.R.S.—Received* 6 *June* 1940.)

This paper describes an investigation of the possibility of measuring small amounts of carbon dioxide in air by the absorption of infra-red radiation. It is shown that the method is simple, trustworthy and accurate, and is very sensitive for small amounts of carbon dioxide, of the order of that present in ordinary air. The principle is applied in a form superior to any hitherto used, in that the whole of the transmitted radiation is measured instead of merely the radiation at the maximum of a particular absorption band. This makes a spectrometer unnecessary, eliminates the disturbing effect of variation of temperature, and allows much less sensitive detecting apparatus to be used. The only preliminary treatment necessary is the removal of water vapour from the air under examination.

The growth of nodule bacteria in the expressed juices from legume roots bearing effective and ineffective nodules. By H. K. Chen, H. Nicol and H. G. Thornton. (*Communicated by Sir John Russell, F.R.S.—Received 6 June* 1940.)

Strains of pea and soy-bean nodule bacteria, differing in their effectiveness in benefiting the host legume, were grown in media containing the unheated root juices from uninoculated host plants and from host plants bearing effective and "ineffective" nodules, and their growth was measured. The growth of the different bacterial strains on root juice from uninoculated plants was not correlated with their effectiveness. The juice from roots with effective nodules produced significantly better growth of the bacteria than juice from roots with ineffective nodules in 27 comparisons out of 44, the differences in the remaining comparisons being insignificant. The juice from roots with effective nodules produced significantly better growth than the juice from uninoculated roots in 10 comparisons out of 25 and significantly poorer growth in 3 comparisons. The juice from roots with ineffective nodules produced significantly poorer growth than the juice from uninoculated plants in 11 comparisons out of 25, and better growth in only 1 comparison.

The production, as a result of infection, of soluble substances affecting growth of the bacteria, affords an explanation of those differences in nodule growth that determine the effectiveness or ineffectiveness of the different strains of bacteria as regards nitrogen fixation within the host plant.

The actions of eserine-like compounds upon frog's nerve-muscle preparations, and conditions in which a single shock can evoke an augmented muscular response. By S. L. Cowan. (*Communicated by E. B. Verney, F.R.S.—Received 15 December* 1939.)

The muscular responses to one shock, or to two shocks separated by 3–10 msec., applied to the nerve of a frog's nerve-sartorius preparation, when immersed in Ringer's solution, containing glucose and buffered with bicarbonate, were reasonably constant for about 3 hr. The main response to two shocks was followed by a small after-tension, which, like Bremer's "neuromuscular contracture", was abolished by atropine in concentration too small to influence the main response, or by curarine in concentrations which somewhat reduced the main response.

Treating nerve-sartorius preparations from frogs which had been kept at 14–18° C for some days before use ('warm' frogs) with prostigmine, eserine, or the dimethylcarbamic ester of 8-hydroxymethylquinolinium methylsulphate usually produced no change in the main response to one or to two shocks, but augmented the after-tension in the response to two shocks, and caused an after-tension to follow the response to a single shock. On the other hand, treating preparations from frogs which had been kept at 0–5° C for more than 40 hr. before use ('cold' frogs) with any of the eserine-like compounds caused the response to one or to two shocks to become augmented, as with mammalian preparations, and to be followed by irregular twitchings which lasted for about a second. The after-twitchings were abolished by atropine in concentrations which slightly reduced the main response, or by curarine in concentrations which reduced the main response rather more.

Treating 'warm' preparations with Ringer's solution of half the usual calcium content, or with Ringer's solution containing any of the following agents (which are known to sensitize to stimuli of long duration), in small concentration, produced no appreciable change in the responses: tetraethylammonium iodide, sodium citrate, guanidine carbonate, or methylguanidine hydrochloride. If, however, preparations so treated were also treated with any of the eserine-like compounds, their responses then became like those of a 'cold' preparation which had been treated with an eserine-like compound.

The responses evoked by applying one shock, or two with an interval of 3–10 msec. between them, to the pelvic end of the frog's sartorius muscle, after sufficient curarine to render nerve stimulation ineffective, were uninfluenced by any of the eserine-like compounds, or any of the five agents mentioned in paragraph 4, provided that shocks of short duration were used; when condenser discharges of long duration were used, treatment of 'warm' preparations with any of the five agents, in the same concentrations as produced the changes described in paragraph 4, augmented the response.

Calcium-deficiency, in large or small degree, affected little the response to stimulation of muscle. But a moderate deficiency changed the responses to repetitive stimulation of nerve in such a way as might have been imitated by stimulating the nerve of the untreated preparation at a lower frequency. When the deficiency was more severe, on stimulating with 150 shocks per sec, no response was visible for the first 1–3 sec., the tension then rose to less than a quarter of its normal maximum, and afterwards fell slowly. Prostigmine in large concentrations produced little improvement, whereas calcium quickly produced a marked improvement.

The actions of eserine-like compounds upon frog's nerve-muscle preparations, and the blocking of neuromuscular conduction. By S. L. Cowan. (*Communicated by E. B. Verney, F.R.S.—Received* 15 *December* 1939.)

The actions of prostigmine, eserine, and the dimethylcarbamic ester of 8-hydroxymethylquinolinium methylsulphate upon the frog's isolated nerve-sartorius preparation have been examined by a method developed by Lucas With Ringer-soaked preparations from frogs kept at 14–18° C for some days before use the minimum interval at which two shocks applied to the nerve could elicit a summated muscular response was about 20 % longer than the absolute refractory period of the nerve. Any of the above-mentioned compounds prolonged the minimum interval for a summated response, but caused the time at which an extra interpolated shock began to cut down the response to the final shock to become only a little later. Curarine or atropine reversed the prolongation of minimum interval.

By the same method, the actions of the same eserine-like compounds upon preparations which had been treated with Ringer's solution of half the usual calcium content were examined. Either before or after treatment, it was impossible to cut down the muscular response by interpolating an extra shock.

The action of prostigmine upon Ringer-soaked preparations was examined by a method developed by Adrian, involving determination of the rate of recovery of excitability in the motor nerve at the site of stimulation and in the remaining more peripheral part of the preparation. Prostigmine influenced little the recovery process at the site of stimulation, whereas it prolonged the slower and more peripheral recovery process. Curarine reversed the prolongation of the more peripheral recovery.

With Ringer-soaked preparations, during the block of neuromuscular transmission produced by rapid repetitive stimulation of the nerve, the response of the muscle to direct stimulation was reduced. If, however, neuromuscular transmission had been blocked by curarine, stimulation of the nerve did not reduce the response of the muscle to direct stimulation.

On the turbulence of a tidal current. By J. PROUDMAN, F.R.S. (*Received 12 June* 1940.)

The paper contains a theoretical investigation of the dynamics of a current when turbulence is of significance. The depth of water is taken to be uniform, while the external body-forces, the mean velocity of the current, and the statistics of turbulence are taken to be uniform over all horizontal planes, though they may vary between the sea-surface and the sea-bottom. The statistics considered are those involved in the forces of internal friction, the rates of doing work and the energy relations. The rotation of the earth and slow variations in time are allowed for.

After deriving a number of formulae which do not appear to have been previously placed on record, though they involve no new principles, a number of particular types of disturbance are studied. They are:

- isotropic turbulence,
- periodic disturbance, neglecting viscosity,
- periodic disturbance, allowing for viscosity.

Only the last of these is able to account for the phenomenon of internal friction. One particular example is worked out, allowing for viscosity, viz. that of a steady current parallel to a vertical plane and uniform between the sea-surface and bottom. In this the forces of internal friction are proportional to a hyperbolic function of the depth below the sea-surface. To obtain a case more like that of actual tidal currents we should require the mean current to be such a function of the depth as would make the forces of internal friction proportional to the depth itself. But the investigation of such a function would require more elaborate mathematical methods than those used in the paper.

The flight of the dipterous fly *Muscina stabulans* Fallén. By F. S. J. Hollick.

(*Communicated by J. Gray, F.R.S.—Received* 17 *June* 1940.)

Females of the dipterous fly *Muscina stabulans* Fallén were chosen for an investigation of flight.

A comparison was made between the flight-system with the insect held stationary and in normal free flight, based on the resultant of those forces which maintain or change its state of motion during flight. Values for this resultant were determined in thirty-two individuals performing regular wing movements when held stationary in 'still air'. In all except four of these, the resultant force generated was equal to or greater than the weight of the insect. On an average, the line of action of the resultant was inclined forwards and upwards at 48° to the body-axis of the insect. The line of action cut the body-axis between 1·75 mm. in front to 3·25 mm. behind the centre of gravity, according to the amplitude of the wing beat. In the majority of cases the point of intersection was behind the centre of gravity in a position incompatible with stable free flight. This last result was confirmed by observing the path travelled by insects performing these wing movements when suddenly liberated in 'still air'. It was found that if the insect is moved forwards at the speeds of normal flight, with the body-axis inclined as in free flight, it continues in the same state of motion on liberation.

The effect of exposing the insect, held stationary, to a stream of air was investigated with the aid of a wind-tunnel. It was found that in those cases where the resultant force passes behind the centre of gravity of the insect in 'still air', the resultant is displaced forwards and comes to pass close behind, through, or in front of the centre of gravity, depending on the velocity of the stream of air and the inclination of the body-axis of the insect. There are good grounds for believing that the living insect held stationary in an appropriate stream of air closely resembles in essentials the system in free flight and is therefore suitable for further study of flight. In analysing further the forward displacement of the line of action of the resultant, it was found that the effect of the stream of air on the body of the fly was negligible in this connexion, as compared with that upon the wings.

By recording photographically the amplitude of wing beat for ten individuals, it was shown that the amplitude decreased with increasing airspeed and inclination of the body-axis—an effect which might be expected to produce a backward displacement of the resultant, that is, in a direction opposite to that observed. A study of the path travelled by the wing relative to the insect revealed a factor that would tend to produce the observed forward displacement of the resultant. It was observed that when the insect is exposed to a stream of air the course of the downward beat is displaced forwards, so as to convert the elliptical track, characteristic of wing movements in 'still air', into the figure '8' path, commonly associated with insect flight. This change in the path travelled by the wings was found to be dependent upon the movement or position of the third antennal joint relative to the second, resulting from the action of the stream of air on the third joint with its arista. The characteristic flight attitude of the legs, and the continued maintenance of wing vibrations when air was blown upon the insect from in front, were also found to be dependent upon the integrity of this system.

The path travelled by the wing relative to the insect is also dependent on the

interaction of the sensory inflow from the antennae and halteres, since after mutilation of the latter the wing-movements associated with stimulation of the former were interrupted.

In the appendices are described: (I) the method of mounting living insects for experiment; (II) the design, calibration and use of an aerodynamic balance suitable for determining the magnitude of the vertical and horizontal components of the 'resultant' and the position of its line of action relative to the centre of gravity of the insect; and (III) a small wind-tunnel, with plate glass top and bottom, for observing the behaviour of the insect when exposed to a stream of air, and for recording photographically the amplitude of wing beat.

The quenching of the resonance radiation of sodium. By R. G. W. NORRISH, F.R.S. and W. MACF. SMITH. (*Received* 17 *June* 1940.)

The effective cross-section for quenching of sodium resonance radiation by the saturated and unsaturated hydrocarbons, tertiary amines and several diatomic molecules has been measured. The results indicate that with regard to quenching ability the gases fall into two groups, one comprising the saturated hydrocarbons and the inert gases and the other the unsaturated hydrocarbons and the amines. The difference in behaviour is too pronounced to be explained in terms of the discrepancy between the amount of energy the sodium atom gives up and the quenching molecule can receive but may be attributed to the presence of unsaturation in the molecule. Within any one series the number of atoms in the molecule apart from the unsaturated centre seems to have little influence on the quenching ability and it has been concluded that the quenching ability may be regarded as proceeding from a centre of unsaturation. The results have been qualitatively considered in the light of Stearn and Eyring's theory of non-adiabatic reactions, and in the terminology of the theory of the intermediate complex we may say that the presence of unsaturation manifests itself in a relatively large transmission coefficient.

On the specific heat of the sodium chloride crystal. By E. W. KELLERMAN. (*Communicated by M. Born, F R S.—Received* 21 *June* 1940.)

Calculations of about 280 proper frequencies of the sodium chloride lattice have been made previously. These frequencies have been used to determine the specific heat of NaCl according to Born's theory. It is shown that this number of frequencies is sufficiently large to make the determination of the frequency distribution and, therefore, of the specific heat possible. Agreement with experiment is obtained. It is shown in agreement with Blackman's investigations that the deviations from Debye's theory as found experimentally are due to Debye's assumption of a ν^2 law for the frequency distribution instead of distribution following from the atomic theory.

ABSTRACTS

OF PAPERS COMMUNICATED TO THE ROYAL SOCIETY OF LONDON

In accordance with a resolution of Council, summaries or abstracts of papers are to be published as soon as practicable. The publication of such abstracts in no way indicates that the papers have been accepted for publication in any fuller form. These abstracts are issued for convenience with the "Proceedings of the Royal Society of London" but do not form a part of the "Proceedings".

---

9 OCTOBER 1940

Catalytic activity, crystal structure and adsorptive properties of evaporated metal films.

By O. BEECK, A. E. SMITH and A. WHEELER. (*Communicated by J. W. McBain, F.R.S.—Received* 4 *April* 1940.)

Metal films of high and reproducible catalytic activity were obtained by condensation of their vapours on glass at any desired temperature. The catalytic activity was measured by the hydrogenation of ethylene. The crystal structure of these films was investigated by electron diffraction. By controlling the pressure of an inert gas (N_2, A, etc.) during evaporation of the metals unoriented and oriented films could be produced at will, and their catalytic activities were compared.

Completely oriented nickel films were obtained with an inert gas pressure of 1 mm., the (110) plane, the least dense of the planes, lying parallel to the backing and the two remaining axes showing random distribution. Iron films were oriented with their (111) plane parallel to the backing, the least dense plane again thus oriented. Low-pressure adsorption of hydrogen at room temperature and of carbon monoxide at liquid air temperature revealed that the oriented gas evaporated nickel films have twice the available surface per gram of randomly oriented high-vacuum films but ten times the activity. The oriented films have therefore fivefold the activity of unoriented films. Oriented films of an available surface equal to unoriented but of fivefold activity could also be obtained in high vacuum by evaporation on to oriented films previously produced by evaporation in an inert gas. The activity per unit weight of the films was constant, indicating ready accessibility to the interior of the film by the reacting gases.

Adsorption of hydrogen was found to be immeasurably fast in all cases. Adsorption isotherms on nickel films were obtained for C_2H_4, CO, N_2, H_2 and O_2; and the effect of catalyst poisoning by CO and O_2, as well as the effect of sintering, was studied.

Loss of activity, decrease of hydrogen adsorption, and amount of poison were found to be proportional.

The most extensive studies were made on nickel films, but films of iron, cobalt, palladium, platinum and copper were also investigated and, with the exception of copper, similar results were obtained. The enhanced activity of oriented films appears to be associated with the larger distances in the (110) plane of nickel or the (111) plane of iron.

On the mechanism of boundary lubrication. Part I. The action of long-chain polar compounds.

By O BEECK, J. W. GIVENS and A. E. SMITH.

(*Communicated by J. W. McBain, F.R.S.—Received* 4 *April* 1940.)

The effect of long-chain polar compounds on the coefficient of kinetic friction under boundary conditions has been studied using the Boerlage four-ball friction apparatus in various modifications. With steel balls of the highest grade, coefficients of friction for a great number of lubricants were measured as a function of the relative velocity of the rubbing surfaces.

The structure of thin films of these lubricants rubbed on polished mild steel surfaces was investigated by electron diffraction. It was found that lubricants showing no or little surface orientation had a constant coefficient of friction of about 0 1 over the available velocity range from 0 to 1 cm./sec. With oils which showed high surface orientation, imparted by addition of long-chain polar compounds, a sudden decrease of the coefficient of the friction was observed at various velocities of the sliding surfaces, depending upon the compound used Investigation of a great number of compounds gave a direct correlation of this effect with molecular orientation: those compounds causing the effect to occur at the lowest velocities were found to be most highly oriented with their carbon chains most nearly perpendicular to the surface Since such a change of the coefficient of friction can only be explained by the wedging of oil under the surface (oil drag), the effect was termed the 'wedging effect' leading to a type of lubrication which may be called 'quasi-hydrodynamic'. By measuring the electrical resistance between the sliding surfaces it was found that the regions of sudden decrease of the coefficient of friction correspond to a change from metallic contact to extremely high resistance.

The investigation shows that long-chain polar compounds act primarily by inducing the wedging effect and not by giving a direct protection to the surface.

On the mechanism of boundary lubrication. Part II. Wear prevention by addition agents.

By O. BEECK, J W. GIVENS and E. C WILLIAMS.

(*Communicated by J. W. McBain, F.R.S.—Received* 4 *April* 1940.)

If two metal surfaces slide over each other in the presence of a lubricant and under high load, high pressures and temperatures prevail at those isolated spots which actually carry the load, leading to wear and possibly to breakdown. The action of wear-preventing agents under these conditions has been studied in detail, and it has been found that such agents are effective through their chemical polishing action, by which the load becomes distributed over a larger surface and local pressures and

temperatures are decreased. Especially effective are compounds containing phosphorus or other elements of group V of the periodic system. These have been found to form a metal phosphide or homologue on the surface which is able to alloy with the metal surface, lowering its melting-point markedly, and by this action aiding greatly in maintaining a polish.

The wear experiments were carried out with a highly sensitive and accurate method which uses metal-plated steel balls as its sliding elements. Under the experimental conditions additions of 1·5 % triphenyl phosphine or triphenyl arsine in white oil gave wear-prevention factors of 7·2 and 12·2 respectively (relative to pure white oil). A further addition of 1 % of a long-chain polar compound is able to double the wear-prevention factor obtained with the polishing agents and wear-prevention factors as high as 17·6 have been observed. The specifically physical action of the long-chain polar compounds is discussed in the preceding paper.

The flame spectrum of carbon monoxide.

By A. G. GAYDON.

(*Communicated by A. C. G. Egerton, Sec.R.S.—Received 13 June 1940*)

The spectrum of the flame of carbon monoxide burning in air and in oxygen at reduced pressure has been photographed on plates of high contrast which display the band spectrum clearly above the continuous background. Greater detail has been obtained than has been recorded previously and new measurements are given.

The structure of the spectrum has been studied systematically. It is shown that the bands occur in pairs with a separation of about 60 cm $^{-1}$, this separation probably being due to the rotational structure. Various wave-number differences are found to occur frequently, and many of the strong bands are arranged in arrays using intervals of 565 and 2065 cm. $^{-1}$.

The possible origin of the spectrum is discussed. The choice of emitter is limited to a polyatomic oxide of carbon, of which dioxide is the most likely. The spectrum of the sub-oxide C_3O_2 shows some resemblance to the flame bands, but this molecule is improbable as the emitter on other grounds. A peroxide CO_3 is also a possibility, but no evidence for the presence of this has been obtained from experiments on the slow combustion of carbon monoxide.

Carbon dioxide in gaseous or liquid form is transparent through the visible and quartz ultra-violet, and the flame bands are not obtained from CO_2 in discharge tubes. Comparison with the Schumann-Runge bands of oxygen shows that it is possible that the flame bands may form part of the absorption band system which is known to exist below 1700 A if there is a big change in shape or size of the molecule in the two electronic states.

The electronic energy levels of CO_2 are discussed. Since normal CO_2 is not built up from normal CO and oxygen, an electronic rearrangement of the CO_2 must occur after the combustion process. Mulliken has suggested that the molecule in the first excited electronic state, corresponding to absorption below 1700 A, may have a triangular form.

The frequencies obtained from the flame bands are compared with the infra-red frequencies of CO_2. The 565 interval may be identified with the transverse vibration ν_2, indicating that the excited electronic state is probably triangular in shape The

2065 interval cannot, however, be identified with the asymmetric vibration ν_3 with any certainty.

If the excited electronic state of CO_2 is triangular then molecules formed during the combustion by transitions from this level to the ground state may be 'vibrationally activated'. This is probably the reason for many of the peculiarities of the combustion of carbon monoxide.

X-ray structure and elastic strains in copper. By S. L. SMITH and W. A. WOOD. (*Communicated by G. W. C. Kaye, F.R.S.—Received* 18 *June* 1940.)

An X-ray examination of copper has been undertaken whilst the metal has been actually under tensile stress, and the X-ray structure investigated at a systematic series of points on the load-extension curve and during unloading and reloading from selected points on the curve. It is shown that the permanent strain is associated with the breakdown of the grains into the crystallite formation and that this change is essentially irreversible. The elastic strain of the metal is accompanied by reversible changes in dimensions of the atomic lattice which take place without leaving in the lattice any permanent distortion, as shown by the observation that the X-ray diffraction rings, including rings of the diffuse type, contract and expand under the action of the applied stress without any change in the degree of diffusion. The lattice changes are distinguished in this way from certain lattice strains or lattice distortions permanently imposed on the lattice as a result of deformation by cold-work. Quantitative measurements are made on the elastic lattice strains exhibited by the (400) and (331) planes in a direction perpendicular to the axis of the applied stress, and these are compared with the equivalent external elastic constants. These measurements show that marked difference in rate of strain may take place in neighbouring grains subjected to the same external stress, and on this difference is based an explanation of the extensive breakdown of the grains into components of widely varying orientations which characterizes the structure of a polycrystalline metal after deformation beyond the yield point.

Infra-red absorption spectra of some amino compounds. By L. KELLNER. (*Communicated by W. T. Astbury, F.R.S.—Received* 25 *June* 1940.)

An investigation of the infra-red absorption spectra of five amino compounds (glycine, diketopiperazine, tetramethyldiketopiperazine, glycylglycine and urea) has been made in the region 2 8–3·6μ. The substances were used in the form of thin crystalline layers deposited on quartz windows.

The spectra are discussed with regard to the molecular structure of the compounds under consideration. It is shown that the number and position of the N—H frequencies in glycine and glycylglycine confirms the view that these two molecules are in the zwitterion form in the crystal. The close similarity between the spectra of diketopiperazine and tetramethyldiketopiperazine on the one hand and the amino acids and urea on the other hand, proves that no lactamlactim interchange occurs in diketopiperazine and its derivative. Both compounds are shown to possess a centre of symmetry. It follows from the experimental evidence that in all investigated substances resonance between the C—N and C═O bonds takes place.

The vibrations and the molecular structure of urea and guanidonium.

By L. KELLNER. (*Communicated by W. T. Astbury, F.R.S.—Received 25 June* 1940.)

The vibrations of urea and guanidonium have been calculated for a field containing valence and angle forces. The assumption is made that urea has the symmetry C_{2v} and guanidonium C_{3h}. It is shown that it is possible to assign every observed frequency of these two substances to definite modes of vibrations under these assumptions. The force constants have been evaluated and have been found to be $f_{\mathrm{C-N}} = 7\cdot1 \times 10^5$ dynes/cm. for guanidonium, and $f_{\mathrm{C-N}} = 6\cdot6 \times 10^5$ dynes/cm and $f_{\mathrm{C=O}} = 9\cdot7 \times 10^5$ dynes/cm. for urea. These values are compatible with the hypothesis that quantum mechanical resonance occurs in both molecules, with the result that the C—N bond in urea has approximately 28% double bond character and the C—O linkage a corresponding single bond character. The guanidonium ion shows complete resonance; each C—N bond has ⅓ double bond character. Curves have been drawn to illustrate the relation between the valence force constants and the bond character.

Liquid jets of annular cross-section.

By A. M. BINNIE and H. B. SQUIRE. (*Communicated by R. V. Southwell, F.R.S.—Received 25 June* 1940.)

Experiments were made on a jet of water issuing from an annular orifice with a low velocity. The effect of surface tension was to cause the jet to coalesce after a short distance. The length of the annular portion was measured, and (when certain precautions were taken) was found to be in fair agreement with the predictions of an approximate theory.

The electrical polarization of frog skin by direct currents.

By O. GATTY. (*Communicated by E. K. Rideal, F.R.S.—Received 27 June* 1940.)

Data concerning the electrical polarization of frog skin by direct currents are presented with adequate statistical control. The results are discussed and show that the skin is electrically non-homogeneous and in a state of disequilibrium. The results are presumptive evidence for at least a three-unit model consisting of a series resistance and local action between an anodic area and a cathodic area as is also considered in the theory of corroding metals.

On stereochemical types and valency groups (Bakerian Lecture).

By N. V. SIDGWICK, F.R.S. and H. M. POWELL. (*Received 27 June* 1940.)

Theory and experiment agree that the arrangements in space of the covalencies of a polyvalent atom, while they are subject, from a variety of causes, to small variations seldom exceeding 5–10°, tend to conform to a quite limited number of types. Recent developments of mathematical physics have made it possible to calculate how these types are related to the electronic groups occupied by the valency electrons.

It is, however, not always possible to say which these sub-groups are, and it is in any case desirable for the chemist that he should be able to relate the spatial grouping to some more familiar property of the atom. The property here used is the size (in G. N. Lewis's sense) of the valency group of the central atom, and the number of shared electrons which it contains, together with that of the preceding (unshared) electronic group in the atom. The experimental results show the following relations.

I. When the valency group is less than 8 we have with a covalency of 2 a linear structure (as in Cl-Hg-Cl): with a covalency of 3 a plane with equal angles of 120° (BF_3)· with 6 electrons of which only 4 are shared, as in the vapour form of stannous chloride, probably a triangular molecule $Sn\langle^{Cl}_{Cl}$.

II. With a complete octet the arrangement can be either tetrahedral or planar. When the covalency is less than 4 (partially shared octet) it is always derived from the tetrahedron, as in the triangular OH_2 and the pyramidal NH_3. The fully shared 4-covalent octet can be of either form. When the preceding electronic group, n in the grouping (n) 8, is 2, 8, or 18, it is always tetrahedral In the transitional elements where $8 < n < 18$ it is found that if n is not much more than 8 the structure is tetrahedral, and if it is not much less than 18, planar; but the two series overlap, and for values of n of 13 to 15 and perhaps 16, both forms occur.

III. When the valency group is 10 with at least 2 of them (the 'inert pair') unshared, the structure of a dicovalent atom (as in $M[I_3]$) is linear, that of the 3-covalent (as in $C_6H_5.ICl_2$) is unknown; while that of the 4-covalent decet is found in the thallous and plumbous compounds to be planar, but in tellurium tetrachloride may possibly be a distorted tetrahedron.

IV. The peculiar 4-covalent duodecet in $M[ICl_4]$ is planar.

V. Covalency 5 So far as it has been examined (4 different types) this has always been found to give a trigonal bipyramid.

VI. Covalency 6. Three structures are theoretically possible: a trigonal prism, a trigonal antiprism, and a regular octahedron. Experimentally the structure is always found to be the octahedron, except in a few giant molecules such as those with the nickel arsenide lattice The regular octahedron has been found with practically every possible size of the preceding group, as well as with the inert pair of electrons.

VII. A covalency of 7 has been examined in three or four compounds, which have two different structures, one derived from an octahedron and the other from the trigonal prism.

VIII. Covalency 8. The only compound examined, $K_4[Mo(CN)_8]$, is found (rather unexpectedly) to have a dodecahedral arrangement of the 8 CN groups.

Nearly (but not quite) all the structures can be even more simply related to the size of the valency group by assuming that the mean positions of the electron pairs in this group are the same whether they are shared or not, the structure being linear for 4 electrons, plane symmetrical for 6, either tetrahedral or plane for 8, a trigonal bipyramid for 10, and an octahedron for 12.

On maintained convective motion in a fluid heated from below. By A. Pellew and R. V. Southwell, F.R.S. (*Received* 1 *July* 1940.)

This paper examines the stability in viscous liquid of a steady regime in which the temperature decreases with a uniform gradient between a lower horizontal surface which is heated and an upper horizontal surface which is cooled. The problem has been treated both experimentally and theoretically by Bénard, Brunt, Jeffreys, Low and Rayleigh, and it is known that instability will occur at some critical value of $gh^3\Delta\rho/\rho k\nu$, h denoting the thickness of the fluid layer, $\Delta\rho/\rho$ the fractional excess of density in the fluid at the top as compared with the fluid at the bottom surface, k the diffusivity and ν the kinematic viscosity. The critical value depends upon the conditions at the top and bottom surfaces, which may be either 'free' or constrained by rigid conducting surfaces.

The theoretical problem is solved under three distinct boundary conditions, and greater generality than heretofore is maintained in regard to the 'cell pattern' which occurs in plan. In addition, an approximate method is described and illustrated, depending on a stationary property akin to that of which Lord Rayleigh made wide application in Vibration Theory.

Mutation effects of ultra-violet light in *Drosophila*. By K. MACKENZIE and H. J. MULLER. (*Communicated by J. B. S. Haldane, F.R.S.—Received 8 July* 1940.)

Experiments are described which show that the frequency of sex-linked lethal mutations induced in spermatozoa of *Drosophila* by ultra-violet irradiation may be raised to approximately 5%. Among the important factors for obtaining such a result are the use of flies with a minimum of pigmentation, exposure through the ventral surface and the use for breeding of those individuals in which, as shown by their earlier death, the radiation has penetrated more effectively. Working by these methods, it has been found possible, by the use of a combination of filters, to determine that the region in which the mutational effect of ultra violet light begins is between λ365 mμ and λ300 mμ, as expected for nucleic acid absorption. Further evidence has been adduced, by the use of an improved genetic technique, substantiating the earlier conclusion of the authors that these rays, although producing gene mutations with the frequency stated, fail to produce gross structural changes (rearrangements) of chromosomes in appreciable quantities. The genetic effect of ultra-violet radiation thus differs decidedly from that of doses of X and higher energy radiation of the same gene-mutational strength. Preliminary data also indicate that minute structural changes (rearrangements) of chromosomes resemble the gross structural changes and differ from gene mutations in being produced with difficulty, if at all, by ultra-violet light, as compared with X-rays. Cases of viable 'terminal deficiencies' also were absent. It seems very probable, therefore, that ultra-violet light cannot, or can only with difficulty, cause 'thoroughgoing' breakage of the chromonema, and that such breakage must occur in the production both of gross and of minute rearrangements but not in the process of gene mutation proper. If the latter process involves breakages at all, these would be of a different kind. It is argued that this evidence of a distinction between gene mutations and structural changes tends to favour the view that the 'genes' correspond to actual segments of the chromonema and are not mere conceptual isolates.

The influence of the solvent on the formation of micelles in colloidal electrolytes. Part I. Electrical conductivities of sodium dodecyl sulphate in ethyl alcohol-water mixtures. By A. F. H. Ward. (*Communicated by E. K. Rideal, F.R.S.—Received* 10 *July* 1940.)

In very dilute solutions in water sodium dodecyl sulphate behaves as a completely dissociated electrolyte. Above a concentration $c = 0{\cdot}00722$ N the equivalent conductance (Λ) falls sharply and micelles are formed. To throw light on the mechanism of micelle formation, the variation of Λ with c has been measured in a series of mixtures of water and ethyl alcohol. The curves for Λ against $\sqrt{c}$ show a gradual transition from the type associated with micelle formation to a uniform curve characteristic of strong electrolytes.

Addition of alcohol decreases the tendency to form micelles When there is 40% or more by weight of alcohol, micelles are no longer formed. The critical concentration first falls on addition of alcohol and then rises again. The effects of alcohol addition can be interpreted satisfactorily by considering the radius of the micelle to be equal to the length of the paraffin chain and thus to be independent of the concentration and nature of the solvent

The energy changes involved in micelle formation are calculated Aggregation can occur if the interfacial energy available from the destruction of the paraffin-solvent interfaces of the ions is greater than the work to be done against electrical repulsion. The interfacial energy is no longer the greater beyond 40% of alcohol It is shown that alcohol molecules are not in solution in the paraffin interior of the micelles in the mixed solvents, but are strongly adsorbed on the micelle surfaces

The influence of regions of lowered dielectric constant around the micelles is considered. This is more important at high concentrations The final fall in Λ with highly supersaturated solutions is attributed to the increased viscosity of the solution.

The high-frequency resistance of superconducting tin. By H. London. (*Communicated by A. M. Tyndall, F.R.S.—Received* 11 *July* 1940.)

The high-frequency resistance of tin in the superconducting state was measured at a wave-length of 20·5 cm. by a calorimetric method based on the principle of eddy current heating. It was found that the resistance decreases *gradually* when the temperature falls below the transition point in contrast to the *sudden* drop in resistance peculiar to direct currents. An explanation of such a behaviour is given, based on the assumption of the simultaneous presence of normal and superconducting electrons. Good agreement between theory and experiment was found.

Absolute measurements of the conductivity in the normal state at low temperatures with both high and low frequencies were carried out and it was found that at the temperature of liquid helium the conductivity for high frequency is considerably lower than for low frequency. This behaviour is possibly due to the fact that the mean free path of the electrons becomes larger than the penetration depth due to skin effect under the conditions of high conductivity and high frequency.

On the structure of multilayers. I. By G. Knott, J. H. Schulman and A. F. Wells. (*Communicated by E. K. Rideal, F.R.S.—Received* 11 *July* 1940.)

X-ray and microscopic examinations of multilayers of certain long-chain esters deposited on transparent bases have been made. An X-ray photograph of a stationary multilayer of ethyl stearate is apparently identical with that obtained by rotating a single crystal of β-ethyl stearate about the long axis of the cell. The spots previously observed on oscillation photographs of multilayers of ethyl stearate are in reality parts of circles, successive arcs of which appear on increasing the angle of incidence of the X-rays on the multilayer. These facts reveal that the multilayer consists of micro-crystals all oriented with one axis in common, this axis being the normal to the surface of the multilayer. The single crystals appear to have grown without interruption throughout the entire thickness of the multilayer, which in these investigations was about 1000 molecules thick.

On the electromagnetic two-body problem. By J. L. Synge (*Communicated by E. T. Whittaker, F.R.S.—Received* 11 *July* 1940.)

This paper is classical in the sense that it does not refer to quantum mechanics, but it is relativistic in the sense of the special theory. It contains a direct attack on the problem of the motion of two charged bodies (hydrogenic atom) on the basis of the following hypotheses. (i) the bodies are particles (mathematical points), (ii) the fields are those given by the retarded potential, (iii) the relativistic equations of motion are obeyed, the force being the Lorentz ponderomotive force without a 'radiation term'. These hypotheses appear to be the simplest available and are logically consistent; the dynamical problem based on them seems worthy of investigation, in particular the question as to whether the motion 'degenerates' to an ultimate collision. Previous approximate investigations by Darwin and Sommerfeld did not reveal degeneracy.

A general method of successive approximations leading to an exact solution is outlined. An invariant E and a 4-vector h_r are defined, E reduces to energy and h_r to angular momentum in the Kepler problem, where the mass-ratio is zero. General expressions for the rates of change of E and h_r are obtained, and formulae developed for their principal parts when the mass ratio is small. If the orbit of the light particle is initially approximately circular it remains so, but its radius decreases steadily to an ultimate collision. This agrees with the classical conclusion that an accelerated charge radiates energy, but the rate is much smaller than that given by the Larmor formula. Except in some final calculations, no assumption of small relative velocity is made, the sole basis of approximation being the smallness of the mass ratio.

The kinetics of the thermal decomposition of acetophenone. By C. N. Hinshelwood, F.R.S. and R. E. Smith. (*Received* 15 *July* 1940.)

The mechanisms involved in the thermal decomposition of aldehydes and ketones are varied, and the relations between them somewhat complex. In particular an interesting contrast in behaviour has recently been found between benzaldehyde and acetaldehyde. An investigation of acetophenone has therefore been made for comparison with acetone.

The thermal decomposition of acetophenone takes place predominantly by the step $C_6H_5COCH_3 \rightarrow C_6H_5CH_3 + CO$, the toluene undergoing a subsequent decomposition to give chiefly benzene, methane and carbon. It differs from that of acetone in yielding hardly any ketene. It is homogeneous and nearly of the first order, with no sharp falling off in rate at 20 mm. There is no retardation by nitric oxide or by greatly increased surface, nor can an increased rate of decomposition be induced by the presence of radicals from decomposing diethyl ether. This is taken as evidence for the absence of reaction chains. In this respect the behaviour resembles that of acetone, but other differences in kinetics exist. These are briefly discussed.

The derivation of mechanics from the law of gravitation in relativity theory. By G. L. CLARK. (*Communicated by Sir Arthur Eddington, F.R.S. —Received* 17 *July* 1940.)

The paper gives a general review of an investigation on some problems on motion in the relativity theory. We begin by discussing the motion of a single particle in a weak gravitational field and obtain both the linear and angular equations of motion; this is followed by brief accounts of the problem of two bodies and that of a rotating rod. In each case considered the equations of motion arise as conditions of integrability of the relativity equations for empty space Only one of these problems has previously been treated by this method, this being the case of two bodies not connected by a material tension. However, this investigation, which was carried out by Einstein, Infeld and Hoffmann, introduces difficult ideas relating to the use of certain spatial surface integrals, in the present paper we avoid the use of these integrals.

An essential feature of the work is that we take a first order solution for the particular problem considered and substitute this solution in the quadratic terms of $G_{\mu\nu}$, and then build up a second order solution. This method can be employed when we discuss the interior of matter, and, as an example, we conclude the paper by investigating the case of a rotating mass of liquid of constant density.

On the Devonian fish *Palaeospondylus gunni* Traquair. By J. A. MOY-THOMAS. (*Communicated by E. S. Goodrich, F.R.S.—Received* 29 *July* 1940.)

The anatomy of *Palaeospondylus* has been redescribed and its affinities rediscussed In the skull the 'hemidomes', 'gammations' and 'Trapezial bars' are shown not to be fused to the neurocranium, but to be visceral articulating elements. The two former are interpreted as parts of the palatoquadrates and the latter as the lower jaws. A hyomandibular and ceratohyal are identified, but the former plays no part in the jaw suspension. The 'rostralia' cannot have formed an anterior ring of cirrhi, because the 'dorsal rostralia' terminate bluntly and may even have formed a continuous plate, and it is concluded that the skull had a prominent rostral prolongation. The presence of paired fins, undoubtedly pelvic and probably also pectoral, is described. The vertebral column and tail previously described upside down are for the first time correctly orientated. The former is shown to have haemal elements throughout its length, and the latter to be heterocercal, the caudal fin being supported by haemal spines and a row of distally bifurcating radials. The view that *Palaeospondylus* has Cyclostome affinities is abandoned on account of the presence of paired fins, centra, a heterocercal tail and jaws *Palaeospondylus* is believed to be related to the Placoderms (Aphetohyoidea) but to represent a new order the Palaeospondyloidea, which in its lack of dermal skeleton and other features suggests a closer affinity between the Placodermi and Elasmobranchii than is generally held to exist.

The equilibrium-diagram of the system silver-zinc.

By K. W. Andrews, H. E. Davies, W. Hume-Rothery, F.R.S. and C. R. Oswin.

(*Received 3 August* 1940.)

The equilibrium diagram of the system silver-zinc has been investigated by thermal, microscopic, and X-ray methods. The liquidus curve agrees with that determined by Haycock and Neville (1897) to a very high degree of accuracy. Interesting observations have been made of the decomposition of the β-phase on quenching. Alloys between 45 and 50 atomic % of zinc can be quenched without any visual signs of decomposition, although the work of Owen and Edmunds shows that some of these alloys readily undergo transformation to ordered body-centred cubic structure when quenched It was found impossible to prevent β-phase alloys containing less than 45 atomic % of zinc from decomposing during quenching, whilst when the zinc content of the β-phase exceeds 50 atomic % of zinc the character of the alloy changes completely, and a number of decomposed or partly decomposed structures are obtained according to the exact conditions of cooling. The $\alpha/\alpha+\beta$-, $\alpha+\beta/\beta$-, and $\beta/\beta+\gamma$-phase boundaries determined in the present work are in general agreement with the results of the earlier workers on this system, and do not confirm the alterations to the diagram recently suggested by Owen and Edmunds who used X-ray methods with quenched filings. The difference is greatest for the $\beta/\beta+\gamma$ boundary and in the present work this has been determined both by microscopic methods and by a new technique involving the use of a high temperature X-ray camera. The remainder of the diagram has been determined in full detail.

The present more accurate determination of the diagram has enabled the following points to be established:

1. The depression of the freezing point of silver by zinc is slightly greater than that produced by equal atomic percentages of indium. This does not confirm the hypothesis of whole number liquidus factors previously suggested by Hume-Rothery, Mabbott and Channel-Evans (1934).

2. The solidus and liquidus curves for the β-phase do not coincide at 50 atomic % of zinc as was proposed by B. G Petrenko (1929). The freezing range of the equi-atomic alloy is approximately 3½°.

3. The determination of phase boundaries by X-ray methods with quenched filings is unreliable for this class of alloy and its limitations are discussed.

The magnetic and other properties of the free electrons in graphite. I

By K. S. Krishnan, F.R.S. and N. Ganguli. (*Received* 16 *August* 1940.)

The graphite crystal has a large free-electron diamagnetism which is directed almost wholly along its hexagonal axis. Over the whole range of temperature over which measurements have been made, namely from 90 to 1270° K, this free-electron diamagnetism of graphite per carbon atom is found to be equal to the Landau diamagnetism per electron of a free-electron gas obeying Fermi-Dirac statistics and having a degeneracy temperature of 520° K.

From this experimental result it is concluded (*a*) that the number of free or mobile electrons in graphite is just one per carbon atom; (*b*) that the effective mass of these electrons for motion in the basal plane is just their actual mass, showing that the movements in this plane are completely free and uninfluenced by the lattice field; (*c*) that on the other hand their effective mass for motion along the normal to the basal plane is enormous, about 190^2 times the actual mass, which indicates that the mobile electrons belonging to any given basal layer of carbon atoms are tightly bound to the layer, though, according to (*b*), they can migrate quite freely over the whole of the layer; (*d*) that this tight binding accounts for the observed low degeneracy temperature of the electron gas in the crystal.

The electron gas in graphite thus conforms to a simple model which is easily amenable to theoretical treatment, and it has a low degeneracy temperature which is conveniently accessible for experiment. It therefore forms a suitable medium for studying the properties of an electron gas.

The conclusions stated in (2) are in accord with the quantal views of the electronic structure of graphite, and also with its Brillouin zones There is one zone which can just accommodate three electrons per atom, and the energy discontinuities at all of its boundary surfaces are large. There is a bigger zone which can just accommodate all the four valency electrons, but the energy discontinuities at those of its faces that are perpendicular to the basal plane are very small.

ABSTRACTS

OF PAPERS COMMUNICATED TO THE ROYAL SOCIETY OF LONDON

In accordance with a resolution of Council, summaries or abstracts of papers are to be published as soon as practicable. The publication of such abstracts in no way indicates that the papers have been accepted for publication in any fuller form. These abstracts are issued for convenience with the "Proceedings of the Royal Society of London" but do not form a part of the "Proceedings".

---

27 NOVEMBER 1940

Ionization, excitation, and chemical reaction in uniform electric fields. IV. A reconsideration of the electron energy balance with special reference to hydrogen.

By R. W. LUNT. *(Communicated by F. G. Donnan, F.R.S.—*

Received 26 *June* 1940.)

This paper attempts to formulate in a general way the balance between the gain and loss of energy by electrons for a system in a steady state comprising an electron (and equal positive ion) cloud drifting through a gas at constant pressure in a uniform electric field and confined within a cylindrical insulating boundary co-axial with the field. Such a system serves as a valuable model for the discussion of excitation and chemical reaction in uniform positive column and other discharges. An attempt is thus made to repair the omissions and correct certain errors which are now seen to exist in an earlier treatment of the problem (Part II, Lunt and Meek 1936). In particular, the earlier model is now seen to fail to correspond with a physically realizable steady state, and the conclusions then reached are consequently suspect. These defects are remedied in the new model by the introduction of processes destroying the various species resulting from collisions between electrons and molecules of the parent gas; this involves a detailed consideration of the composition of the gas, especially with regard to the concentration of these secondary species. In illustration, the numerical estimation of the electron energy losses in the principal collision processes in hydrogen is reconsidered; it is found that the earlier conclusions can be supported but in a more precise and extended form.

Ionization, excitation, and chemical reaction in uniform electric fields. V. A reconsideration of Kirkby's investigation of the uniform positive column reaction in electrolytic gas. By R. W. Lunt, C. A. Meek and G. E. Swindell. (*Communicated by F. G. Donnan, F.R.S.—Received* 26 *June* 1940.)

After reviewing the experimental observations and the theory of Kirkby, an attempt is made to assess the more detailed quantitative theory of excitation as effected by an electron swarm drifting in a uniform electric field (which is developed in earlier papers of this series) by considering how far its application to the particular conditions in Kirkby's experiments provides an explanation for his observations. The present theory of such excitation is a more precise analysis based on Kirkby's ideas, but in its application a number of new considerations are taken into account

It is found that the present theory provides a detailed explanation for the large effect exerted on the rate of reaction by the ratio of the electric field in the positive column to the gas pressure, which effect is Kirkby's main discovery. But it also provides a detailed explanation for certain differences in the rate of reaction between one experiment and another which are revealed by a fresh scrutiny of Kirkby's data, and which he attributed to random experimental errors: the origin of these differences is now found to be ascribable to effects depending on the particular values of the current and the gas pressure selected in the individual experiments.

Furthermore, this detailed explanation of Kirkby's observations is consistent with a number of largely independent experimental data and theoretical considerations, notably: the experimental and theoretical data for single collisions between electrons and hydrogen, oxygen and other molecules, the experimental data for the electron motion in other positive column discharges, the experimental data for the excitation of the continuous spectrum and for the generation of atoms in positive column discharges in hydrogen, the experimental data for the synthesis of ozone in discharges in oxygen, the experimental data and the detailed explanation for the negative glow reaction in electrolytic gas, and the kinetics of the oxidation of hydrogen as based on the experimental data for the reaction in hydrogen-oxygen mixtures as initiated by light absorption.

Wind-tunnel correction for a circular open jet tunnel with a reflexion plate. By B. Davison and L. Rosenhead, D.Sc. (*Communicated by J. Proudman, F.R.S.—Received* 19 *July* 1940.)

It is advantageous from many points of view to make test-models as large as possible. One method of doing this is to measure the characteristics of half the model in existing wind-tunnels. One half of the aerofoil is mounted horizontally on a vertical reflexion plate and the plate is placed in a suitable position in an open jet which, in the undisturbed state, is of circular section. The contour of the jet is distorted, especially with models of large semi-span, but this distortion is neglected in the analysis. The correcting factor associated with 'uniform' distribution of lift is worked out exactly and that associated with 'elliptic' distribution approximately.

The effect of the induced downwash on the distribution of lift is ignored. The results are given in suitable tables and figures.

Throughout the working range of normal experiments the correcting factor is of the same order of magnitude as that obtaining when a full model is tested in a jet of circular section.

The asymptotic expansion of integral functions. By E. M. WRIGHT.

(*Communicated by G. H. Hardy, F.R.S.—Received 27 July* 1940.)

1. In a former paper I deduced the asymptotic expansion of the integral function

$$f(x) = \sum_{n=0}^{\infty} \frac{\phi(n)\, x^n}{\Gamma(\kappa n + \beta)} \quad (R(\kappa) > 0)$$

for large x from asymptotic properties of the function $\phi(t)$. In particular, $\phi(t)$ had to be regular and its asymptotic behaviour had to satisfy a certain 'condition A' throughout the half-plane $R(\kappa t) > K$. My results included as special cases most of the known results about the asymptotic expansion of integral functions.

The next step is to widen our class of functions further by relaxing the conditions on $\phi(t)$. I here suppose only that $\phi(t)$ is regular and satisfies condition A in a region of the t-plane including the positive half of the real axis and lying within the half-plane $R(\kappa t) > K$; this region is the 'sector' in which

$$-\mu_1 \leqslant \arg(\kappa t) \leqslant \mu_2, \quad |t| > K, \tag{1}$$

where μ_1, μ_2 are any real numbers satisfying

$$-\tfrac{1}{2}\pi < -\mu_1 \leqslant \gamma \leqslant \mu_2 < \tfrac{1}{2}\pi$$

and $\gamma = \arg \kappa$ From this I deduce the asymptotic expansion of $f(x)$ in that part of the x-plane in which

$$-\min(\mu_1, \omega_2) \leqslant \arg x^{1/\kappa} \leqslant \min(\mu_2, \omega_1), \tag{2}$$

where $\omega_1 = \omega(\mu_1, -\gamma)$, $\omega_2 = \omega(\mu_2, \gamma)$ and $\omega(\mu, \gamma)$ is a real transcendental function of μ, γ and $|\kappa|$. The relevant properties of the function ω are investigated.

The region (2) is bounded by two spirals (degenerating to straight lines if κ is real). If $\mu_2 = \gamma$ or $\mu_1 = \gamma$ the two spirals coincide and the region (2) consists only of the resulting spiral Hence if $\phi(t)$ is regular and satisfies condition A on the positive half of the real axis, we can find the asymptotic expansion of $f(x)$ on the spiral $\arg x^{1/\kappa} = \gamma$.

If $R(1/\kappa) < \tfrac{1}{2}$, a number $\mu_0 = \mu_0(\gamma)$ exists such that, if $\mu_2 = \mu_0$ and $\mu_1 = 2\pi R(1/\kappa) - \mu_0$, the region (2) consists of the whole plane and we have the expansion of $f(x)$ for all large x Since the sector (1) still lies within the half-plane I deduce the results given in a former paper when $R(1/\kappa) < \tfrac{1}{2}$ from hypotheses substantially less than those of that paper.

I also show that my results are now, in an obvious sense, best possible; that is, if the conditions stated in my theorems are further relaxed, the conclusions are false.

The crystal structure of Rochelle salt (sodium potassium tartrate tetrahydrate). By C. A. Beevers and W. Hughes. (*Communicated by W. L. Bragg, F.R.S.—Received* 1 *August* 1940.)

The complete crystal structure of Rochelle salt (sodium potassium tartrate tetrahydrate) has been determined by Fourier and Patterson methods. Some of the difficulties in the application of these methods are discussed.

The tartrate molecule is found to lie approximately in three planes, the planes of each half of the molecule being inclined at 60° to the plane of the carbon atoms. The tartrate molecules are bonded to sodium and potassium atoms both directly and through the medium of water molecules. If the water molecules are to preserve their customary tetrahedral 'bonding' it is necessary to suppose that one of the carboxyl groups of the molecule is also a dipole. A reversal of the continuous chain of carboxyl-water-water dipoles is a possible explanation of the peculiar dielectric properties of the salt.

The effects of low temperature and acclimatization on the respiratory activity and survival of ram spermatozoa. By Min-Chueh Chang and Arthur Walton. (*Communicated by J. Hammond, F.R.S.—Received* 17 *August* 1940.)

The respiratory activity of ram spermatozoa was measured before and after experimental treatment at low temperature. Sudden cooling has a harmful effect (temperature shock) on the subsequent respiratory activity of the spermatozoa. Temperature shock increases as the temperature is lowered. Temperature shock increases with the time of exposure to the lower temperature. When semen is cooled gradually in stages of 5° C and with an interval between stages, acclimatization of the spermatozoon increases with the length of the interval If the interval between stages is 2 hr., the semen can be cooled to 1° C and stored at this temperature for 24 hr. without appreciably affecting the subsequent respiratory activity. Rapid warming has no harmful effect upon the subsequent respiratory activity of the spermatozoa. Temperature shock and acclimatization are factors which affect the apparent optimum temperature for the storage of spermatozoa With rapid cooling the optimum is higher than with slow cooling. Semen cooled gradually to 1° C and stored at this temperature for 6 days retained 74 % of its initial respiratory activity.

The production of cancer by pure hydrocarbons. V. By G. M. Badger, J. W. Cook, F.R.S., C. L. Hewett, E. L. Kennaway, F.R.S., N. M. Kennaway, R. H. Martin and A M. Robinson. (*Received* 20 *August* 1940.)

Tests for carcinogenic activity have been carried out with some 70 compounds, mostly new substances specially synthesized for the purpose, and related in molecular structure to the known carcinogenic compounds. The results support the view that there is a definite association between carcinogenic activity and molecular structure although the exact limits of structural conditions necessary for carcinogenic activity cannot be precisely defined. Among the compounds related to 1 : 2-benzanthracene and 3 : 4-benzphenanthrene, which comprise the majority of the compounds examined, it is evident that there are certain positions of the molecule in which substitution is especially favourable for carcinogenic properties. In the 3 : 4-benzphenanthrene series, which has not hitherto been extensively investigated, positions 1 and 2 seem to be the favourable positions of substitution.

Comparative tests carried out with stock mice have shown that the potency of a compound in the production of epitheliomas in mice does not always run parallel to its potency in the production of sarcomas by subcutaneous injection. Many compounds have been tested by both modes of administration, and in some cases the power to produce epithelioma was markedly superior to the sarcoma-producing potency, whereas in other cases the reverse was true.

A preliminary study has been made of the incidence of tumours of the lung and of the stomach in mice receiving carcinogenic compounds by the aforementioned two and in some cases a third (oral) method of administration. Very suggestive indications have been found that some of the compounds (the 5-*n*-butyl, 5-*n*-amyl, and 5-*n*-hexyl derivatives in the 1 : 2-benzanthracene series, and the 1- and 2-substituted derivatives as opposed to the 6-, 7- and 8-substituted compounds in the 3 : 4-benzphenanthrene series) are peculiarly effectual in producing tumours of the lung and the stomach. These same groups of compounds are also superior in their capacity to produce epithelial tumours by contrast with their lack of effect on the subcutaneous tissues, and the same is true of the three dibenzfluorenes which have been examined.

On the Bicoecidae: a family of colourless flagellates. By L. E. R. Picken. (*Communicated by J. Gray, F R S.—Received* 24 *August* 1940.)

Two distinct species appear to have been confused by James-Clark in his original description of *Bicoeca lacustris* The two forms have been re-examined and photographed, and it is suggested that the differences between them justify their separation into two species, as suggested by Štolc, the one form to retain the name *B. lacustris*, the other to receive that of *B. vacillans* Štolc.

A new species, *B. maris*, is described, and the affinities of the Bicoecidae, Bodonidae and Craspedomonadidae are discussed.

Molecular structure in relation to oestrogenic activity: derivatives of 4 : 4′-dihydroxydiphenylmethane. By N. R. Campbell. (*Communicated by Sir Robert Robinson, F.R.S.—Received* 26 *August* 1940.)

The oestrogenic activity of four classes of 4 : 4′-dihydroxydiphenylmethane derivatives has been studied. Variation of activity in homologous series has been found, and maximum potencies established The relationships between structure and maximum potency have been found parallel to those in other series of synthetic oestrogens.

The effect of methyl groups attached to the aromatic rings has been investigated and the results suggest a biological connexion between these groups and groups on the central carbon atom, since the effects vary in marked manner with alteration of the chain length of the central groups.

The apparent sizes of atoms in metallic crystals with special reference to aluminium and indium and the electronic state of magnesium. By W. Hume-Rothery, F.R.S., and G. V. Raynor. (*Received* 4 *September* 1940.)

The interatomic distances in crystals of alloys cannot be accounted for by assigning a fixed atomic radius to each kind of atom, and the causes of this variation are discussed with special references to the Brillouin zone characteristics of different structures. According to the theory of H. Jones, the effect of an overlap across the side of a Brillouin zone is to compress the zone at right angles to the face concerned, and so to expand the crystal lattice in the same direction. This expansion is not a property of an atom which can be transferred to any other structure, but is a characteristic of a structure with sufficient electrons to produce an overlap. The lattice spacings of alloys of aluminium and indium with copper, silver, gold, and magnesium are examined, and the apparent sizes of the aluminium and indium atoms are discussed, and are shown to be in agreement with the theory. The previous suggestion that in metallic aluminium, the atoms exist in an incompletely ionized state is improbable, and is no longer required in order to explain the facts. New experimental data for the lattice spacings of solid solutions of aluminium and indium are presented, and these show that, whilst the curves connecting the a parameter with the composition are smooth and continuous, the corresponding curves for the c parameter show an abrupt change in direction at about 0·75 atomic % of indium or aluminium. This is taken to imply that, although in metallic magnesium with 2 electrons per atom, the overlap of the first Brillouin zone is in the a direction only, the structure is so near to the stage at which the c overlap sets in that the addition of less than 1 electron per 100 atoms causes the c overlap to take place.

The structure of melamine $C_3N_6H_6$. By I. E. Knaggs and K. Lonsdale. With a note on the optical properties of melamine by G. Williams and R. G. Wood. (*Communicated by Sir William Bragg, P.R.S.—Received* 5 *September* 1940.)

Magnetic and X-ray measurements on crystalline melamine, $C_3N_6H_6$, indicate that the molecular formula is of the amide type, based on the cyanuric ring, but that weak hydrogen bridges exist between the NH_2 groups of 1 molecule and the ring nitrogens of others, so that the substance may also tend to behave as an imide. The density is 1·571, the space-group $P\,2_1/a$, and there are 4 molecules in the unit cell. An approximate structure is given. New optical data confirm that this structure is of a layer type, in which respect it is similar to all other cyanuric ring compounds so far examined.

The mobility of positive ions in their own gas. By R. J. Munson and A. M. Tyndall, F.R.S.—(*Received* 9 *September* 1940.)

Values for the mobility of positive ions of neon in neon, argon in argon, krypton in krypton and xenon in xenon have been obtained in pure gas. It is possible to deduce the mobility of an ion of the same mass as the gas atom from the known values for alkali ions in the gas on the assumption that the same mass relationship holds. In each case the observed value is considerably less, the ratio of the two being of the order of 0·60–0·75. This result may be attributed to the phenomenon of electron exchange in the case of ions in their own gas which does not arise with alkali ions.

The variation of mobility with field E and pressure p has been studied. In Kr and Xe there is a marked fall of mobility with increase of E/p, at higher values the velocity of the ion increasing in a linear manner with increase of $(E/p)^{\frac{1}{2}}$.

The mobility of alkali ions in gases. IV. Measurements in gaseous mixtures. By H. G. DAVID and R. J. MUNSON. (*Communicated by A. M. Tyndall, F.R.S.—Received* 9 *September* 1940.)

The mobilities of some of the alkali ions have been measured in binary mixtures of the inert gases. Blanc's relation that the reciprocal of the mobility is a linear function of the concentration of one of the constituents is accurately obeyed in all cases except for Li^+ in He-Xe mixtures, where the maximum deviation is 4 %.

With clustered Li^+ ions in mixtures of water vapour and an inert gas, deviations from this simple relation are observed and increase in magnitude progressively with the density of the gas. These deviations, due to the polar nature of the water molecule, can be explained by assuming a gradual increase in the size of the clustered ion with increase in the percentage of water vapour.

The mobility of alkali ions in gases. V. Temperature measurements in the inert gases. By K. HOSELITZ. (*Communicated by A. M. Tyndall, F.R.S.—Received* 9 *September* 1940.)

The variation with temperature of the mobility of Li^+ ions in helium, K^+ ions in argon, Rb^+ ions in krypton and Cs^+ ions in xenon has been measured over a considerable range of temperature. Comparison of the results with existing classical theories suggests that the approximate agreement found by Pearce over a similar range of temperature for Cs^+ and Na^+ in He does not hold in general. A quantum mechanical treatment for each individual case seems to be required and for this purpose the experimental data here recorded should be of value.

The distribution of electricity in thunderclouds. II. By G. D. ROBINSON and Sir GEORGE SIMPSON, F.R.S. (*Received* 10 *September* 1940.)

The investigation of the distribution of electricity in thunderclouds described by Simpson and Scrase in 1937 has been continued at Kew Observatory. Alti-electrographs, which record the sign of the potential gradient, are sent up on small free balloons. Additional observations made during eight thunderstorms in the years 1937, 1938 and 1939 are discussed and the conclusions reached by Simpson and Scrase are confirmed. Each thundercloud has positive electricity in the upper half of the cloud, negative electricity in the lower half and in most storms if not in all there is a concentrated positive charge below the main negative charge. The generation of the positive and negative charges in the main body of the cloud is ascribed to the impact of ice crystals and that of the positive electricity in the base of the cloud to the breaking of rain drops in an ascending current of air

Exchange effects in the theory of the continuous absorption of light. I. By D. R. BATES and H. S. W. MASSEY, F.R.S. (*Received* 10 *September* 1940.)

The continuous absorption coefficients of normal Ca and Ca^+ are calculated. Both the discrete and continuous wave functions used in the calculation include the effect of electron exchange. Comparison is affected with results obtained neglecting exchange and it is found that appreciable modifications are introduced by its inclusion. This is particularly true for Ca^+ which is a sensitive case owing to very strong interference in the integrand of the transition matrix element. The bearing of the results on the calculation of absorption coefficients in general is discussed and it is pointed out that the discrepancy between theory and experiment for potassium arises because this also is a very sensitive case.

An application of the results for Ca is made to resolve a discrepancy between determinations of interstellar electron densities carried out using different methods by Stromgren and Struve.

The polarization of electrons by double scattering. By H. S. W. MASSEY, F.R.S., and C. B. O. MOHR. (*Received* 10 *September* 1910.)

The scattering of electrons by gold, xenon and krypton atoms has been investigated, using Dirac's equations The polarization to be expected by double scattering at 90° has been studied for an energy range from 100 to 150,000 eV. The results agree substantially in the energy range 10,000–150,000 eV with those obtained by Mott for an unscreened gold nucleus. It is found, however, that for lower energies the effect of screening is more important. In particular at energies for which the intensity of single scattering at 90° is near a minimum, a large polarization is to be expected in the case of gold. For xenon the polarization never exceeds 4 % and for krypton 2 %.

Modifications of the interaction between a nucleus and electron which would reduce the polarization are considered.

As an example of a field as different as possible from the Coulomb type, the polarization produced in double scattering by a potential well is investigated in detail.

Application of the coincidence method for measurements of short life periods. By JÓZEF ROTBLAT. (*Communicated by J. Chadwick, F.R.S.—Received* 16 *September* 1940.)

A method is described for the determination of short radioactive and nuclear life periods. It consists in varying the length of the electrical impulses produced in a coincidence arrangement of two Geiger-Muller counters. The method can be used for the measurement of life periods in the range 10^{-7} to 10^{-1} sec. The half-life period of radium C′ has been measured in this way and found to be $1{\cdot}45(\pm 0{\cdot}05)\times 10^{-4}$ sec. Experimental evidence has been obtained for the absence of γ-rays in the transition radium C′ → radium D, and some conclusions regarding the lifetimes of excitation states of radium C′ have been drawn The determination of the intensity of weak sources and of the efficiency of Geiger counters for various types of radiation is described. The effect of a time lag in Geiger counters on the measurement of coincidences is discussed.

The respiratory function of the haemoglobin of the earthworm. By M. L. Johnson. (*Communicated by H. Munro Fox, F.R.S.—Received* 17 *September* 1940.)

The oxygen consumption of earthworms (*Lumbricus herculeus* Savigny) has been measured at 10° C, in the dark, in atmospheres containing 20, 10, 5, 2 5 and 1 % of oxygen (i.e. at partial pressures of oxygen of about 152, 76, 38, 19 and 8 mm mercury) with and without the addition of enough carbon monoxide to saturate the haemoglobin of the blood.

In the absence of carbon monoxide the rate of oxygen consumption was the same at 152 and 76 mm.; below 76 mm. it fell sharply.

The rate of oxygen consumption of carbon monoxide-treated worms was lower than that of normal worms at oxygen pressures of 152, 76, 38 and 19 mm. This lowering of rate was not due to the effect of carbon monoxide on respiratory enzymes, for the oxygen consumption of sliced earthworm tissue was not lowered by the amounts of carbon monoxide used in the experiments on whole worms. It is concluded therefore that haemoglobin acts as an oxygen carrier at atmospheric partial pressure of oxygen and at lower oxygen pressures down to 19 mm.

The amount of oxygen carried by the haemoglobin was probably greatest when the pressure of oxygen was about 76 mm. mercury. It is deduced that the loading pressure of earthworm haemoglobin may be as high as 76 mm., and is certainly higher than 19 mm.

The haemoglobin of the blood was responsible for supplying about 20 % of the respired oxygen when the oxygen pressure was at 152 mm , 35 % at 76 mm , 40 % at 38 mm., and 20 % at 19 mm. of oxygen.

The characteristics of thermal diffusion. By S. Chapman, F.R.S. (*Received* 26 *September* 1940.)

When a gas-mixture is contained in a vessel in which a steady temperature gradient is maintained, a concentration gradient is in general set up, whose amount is determined by the logarithm of the temperature ratio, and by k_T, the thermal diffusion ratio; the general theory of non-uniform gases gives successive approximations to k_T, and the first of these, $[k_T]_1$, is accurate within a few per cent. The paper discusses the dependence of $[k_T]_1$ on (*a*) the ratio of the molecular masses; (*b*) their concentration ratio (c_1 or c_2); (*c*) the two ratios of the molecular diameters, inferred from the coefficient of viscosity, to their joint diameter, inferred from the coefficient of diffusion; and (*d*) three parameters depending on the mode of interaction between the unlike molecules. When this interaction is according to the inverse-power law, the three parameters (*d*) are all expressible in terms of the mutual force-index, and $[k_T]_1$ is a function of five independent variables. The general nature of its dependence on these variables is discussed, with particular reference to the end values (for c_1

or c_2 zero) of the thermal diffusion factor α, given by k_T/c_1c_2; these end values involve fewer variables (less by two) than the general values, and their functional character can be represented graphically. It is shown that k_T may be zero not only when c_1 or c_2 is zero, but also for at most one intermediate mixture-ratio. Formulae for $[k_T]_1$ appropriate to various special cases are also given.

An investigation of the disintegration of boron by slow neutrons.

By R. S. Wilson. (*Communicated by J. Chadwick, F.R.S.—Received 27 September* 1940.)

The disintegration of boron by slow neutrons has been investigated using an ionization chamber filled with boron trichloride in conjunction with a linear amplifier.

The magnitudes of the ionization impulses, measured by the oscillograph deflexions, should give under appropriate conditions a measure of the energy release in the disintegration process. The various factors which can affect the size of the oscillograph deflexion are discussed so as to define the appropriate conditions of experiment.

Evidence is found for two disintegration energies. Assuming that the greater energy release corresponds to the formation of the Li^7 nucleus in the ground state and is therefore 2·99 MeV, then the smaller energy, which is released in about 93–94 % of the disintegrations, is 2·57 ± 0·05 MeV. An explanation is offered of the contradiction with the results of Maurer and Fisk. The γ-radiations associated with the reaction have been detected and a rough measurement of the quantum energy has been made.

Molecular anisotropy of urea and of related compounds.

By K. Lonsdale. (*Communicated by Sir William Bragg, P.R.S.—Received 4 October* 1940.)

The magnetic anisotropy of urea has been measured and compared with that of the carbonate and nitrate groups. It does not appear that the hydrogen bonds contribute appreciably to the anisotropy; the latter indicates about the same degree of resonance in all three groups. The anisotropy of the refractivities is greater in urea than in the nitrate group, and considerably greater than in the carbonate group, showing that the $(NH_2)^-$ radical must be much more polarizable than the O^- ion, allowance being made for the interionic distances. Urea nitrate, urea oxalate, cyanuric acid dihydrate and anhydrous cyanuric acid all crystallize in layer structures, as shown by cleavage, X-ray data, magnetic and optical observations. The magnetic anisotropy of urea nitrate and urea oxalate is approximately additive, but that of cyanuric acid dihydrate is considerably larger than the additive value for three urea molecules, being comparable with that of other cyanuric ring compounds. Although the molecules in each layer must be joined together by hydrogen bonds, it does not follow that hydrogen bonds form the links between successive layers, since the layer to layer distance of 3·05 ± 0·10 A is characteristic even of cyanuric triazide, which contains no hydrogen bonds.

Submarine seismic investigations. By E. C. BULLARD and T. F. GASKELL. (*Communicated by Sir Gerald Lenox-Conyngham, F.R.S.—Received* 9 *October* 1940.)

The refraction seismic method has been used to investigate the form of the surface of the hard rocks underlying the sediments on the continental shelf to the west of the English Channel. This surface is found to slope steadily downwards on receding from the land, and to reach a depth of over 8000 ft. at the 100 fm. line. The velocity of elastic waves in the sediments is about 6000 ft./sec. near the surface and up to 9700 ft./sec. lower down; compared to 16,000–22,000 ft./sec. in the basement. The bearing of these results on the structure and history of the shelf is discussed.

The theory of the divisions in Saturn's rings. By G. R. GOLDSBROUGH, F.R.S. (*Received* 9 *October* 1940.)

In an earlier paper on this subject the author (1921) proposed a theory of the rings which showed satisfactory agreement with the observed measurements of the rings. The mathematical method was, however, subjected to criticism. In the present paper the subject is again attacked by an entirely different method which is free from the objections raised against the first method.

A family of periodic orbits of the particles forming the ring, when perturbed by a satellite, is constructed, and the stability of these orbits is examined by the method of small displacements. Stability determined in this way is shown to have a real meaning when applied to the problem in hand.

The positions of instability of the particles lead to the divisions in the ring and the inner and outer boundaries, close agreement with observation being obtained. The analysis, though quite different from that of the earlier paper, reproduces its main features, and introduces further points of interest.

Observations on the chemical constitution of galactogen from *Helix pomatia*: *l*-galactose as a component of a polysaccharide of animal origin. By E. BALDWIN and D. J. BELL. (*Communicated by Sir F. Gowland Hopkins, O.M., F.R.S.—Received* 15 *October* 1940.)

The polysaccharide galactogen, isolated from the albumin gland of *Helix pomatia*, has been shown to be built up from both *d* and *l* isomers of galactose in the probable proportions of 6:1. Crystalline derivatives of *dl* galactose have been obtained from the polysaccharide, thus affording what is believed to be the first proof of the presence of *l*-galactose components in a substance of animal origin. The products of methanolysis of methylated galactogen have been analysed; they are composed of tetramethyl methylgalactosides and dimethyl methylgalactosides in molecular proportions closely agreeing with the simple ratio of 4:3. The former material consists of a mixture of one molecule of 2:3:4:6-tetramethyl methyl-*l*-galactosides with three molecules of 2:3:4:6-tetramethyl methyl-*d*-galactosides, while the dimethyl methylgalactosides consist entirely of the 2:4-dimethyl-*d*-derivatives. Possible structures for the chemical molecule of galactogen are discussed.

The fine structures of phloem fibres. II. Untreated and swollen jute. By R. D. Preston. (*Communicated by W. T. Astbury, F.R.S.—Received* 16 *October* 1940.)

Studies of the jute fibre by X-ray methods, under the polarizing microscope, and after swelling with various reagents, lead to the conclusion that the optical heterogeneity of the wall is not to be explained in terms of changes in cellulose chain direction. The evidence available shows that the wall is composed of chains forming a single, steep spiral. The wall spiral becomes steeper as the cell elongates and this continues until the onset of wall thickening.

The fibre in many ways provides a study complementary to that of hemp. The outer layers of the wall are the least lignified and have the least resistance to swelling. The phenomenon of 'ballooning' is therefore never observed, though structures resembling balloons may be produced by a special technique. For similar reasons transverse cracks are never observed during the swelling process.

Impregnation with lignin seems to cause an increase in thickness of the wall even in the outer layers which are far removed from the cytoplasm-wall interface

The results fully confirm the conclusions already reached that observations of swollen material are usually entirely misleading.

Radio echoes and cosmic ray showers. By P. M. S. Blackett, F.R.S., and A. C. B. Lovell. (*Received* 22 *October* 1940.)

It is suggested that the origin of some of the transient ionic clouds, generally assumed to be responsible for the low level sporadic radio reflexions, may be due to large cosmic ray showers.

It is shown that cascade cosmic ray showers of sufficient energy to produce some of these radio reflexions certainly exist, but there is insufficient published evidence to decide whether any of the echoes already observed are actually due to such showers. More conclusive evidence could be obtained from the frequency-size distribution of the radio echoes observed from a horizontal or vertically directed beam.

I A R I 75.

INDIAN AGRICULTURAL RESEARCH
INSTITUTE LIBRARY
NEW DELHI.

Date of issue.	Date of issue.	Date of issue
6.7.65		

S. C. P.—1/8/47 P. J.-3-5-48-2000